DOCUMENT RESUME

ED 047 486 | 40 | EC 031 910

AUTHOR Cook, John J.; Blessing, Kenneth R.
TITLE Class Size and Teacher Aides as Factors in the Achievement of the Educable Mentally Retarded. Final Report.
INSTITUTION Wisconsin State Dept. of Public Instruction, Madison. Div. for Handicapped Children.
SPONS AGENCY Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
BUREAU NO BR-6-2620
GRANT OEG-3-6-062620-1873
NOTE 326p.

EDRS PRICE EDRS Price MF-$0.65 HC-$13.16
DESCRIPTORS Behavior Rating Scales, *Class Size, *Educable Mentally Handicapped, *Exceptional Child Research, Intermediate Grades, Mentally Handicapped, Primary Grades, *Student Behavior, *Teacher Aides, Teacher Attitudes

ABSTRACT

The purpose of the study was to determine the effects of class size and the use of teacher aides on pupil behavior and teacher behavior in educable mentally handicapped classes at the primary and intermediate grade levels. In the course of the study 20 classes remained constant and 10 aides were employed throughout the project. Effects on pupil behavior were determined by contrasting the educational and linguistic achievement and social behaviors in the several instructional arrangements. Researchers concluded that the major impact of the aides alone appeared to be a reduction in undesirable behavior in the classroom and an enhancement of the spelling achievement; class size alone had an effect on arithmetic and classroom behavior, with the standard classes gaining more in math and manifesting less negative behavior than the increased classes. It was generally felt that the most pronounced effects of the study appeared in the primary classes and suggested to the researchers that administrative strategies other than aides and class size should be developed to enhance the behavioral, linguistic functioning of intermediate educable mentally handicapped pupils. (CD)

EC031018

OE/BEH BR 6-2620
PA 40
EC

ED047484

FINAL REPORT

Project No. 6-2620
Grant No. OEG-3-6-062620-1679

CLASS SIZE AND TEACHER AIDES AS FACTORS IN THE ACHIEVEMENT OF THE EDUCABLE MENTALLY RETARDED

Wisconsin Department of Public Instruction
William C. Kahl, Superintendent
Division for Handicapped Children
John W. Melcher, Administrator
126 Langdon Street
Madison, Wisconsin 53702

August 1970

EC 031 916

Department of Health, Education, and Welfare

U.S. Office of Education
Bureau of Education for the Handicapped

EC031918

2

ED047484

FINAL REPORT

Project No. 6-2828
Grant No. OEG-3-6-062828-1879

CLASS SIZE AND TEACHER AIDES AS FACTORS IN THE ACHIEVEMENT OF THE EDUCABLE MENTALLY RETARDED

John J. Cook, Ph.D.
Kenneth R. Blessing, Ph.D.
Wisconsin Department of Public Instruction
Division for Handicapped Children
126 Langdon Street
Madison, Wisconsin 53702

August 1970

The research reported herein was performed pursuant to a grant with the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Department of Health, Education, and Welfare

U.S.Office of Education
Bureau of Education for the Handicapped

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY

CONTENTS

PAGE

List of Tables vi

List of Figures viii

Acknowledgments ix

Summary xi

Preliminary Comments xii

Chapter I: Problems Under Consideration 1

A. Teacher aides and class size as factors in pupil achievement and classroom behavior 1

B. A behavior rating scale for the retarded 2

C. An approach to language analysis in the retarded 2

D. Filming in the Classroom 2

E. Influence of meteorological factors on behavior in the classroom 3

F. Teacher expectancy and achievement in the mentally retarded 3

G. A look at some assumptions in special education 4

H. The utility of MA as a predictor of achievement 4

I. Cost analysis 4

J. Other considerations 5

Chapter II: Review of the Literature 6

Chapter III: Design and Analysis Considerations 16

PAGE

Chapter IV: Method 27

A. Selection and characteristics of the sample 28

B. Dependent measures 28

Chapter V: Results and Discussion 33

A. Quantitative 33

1. Analyses of dependent measures 33

2. Automatic interaction detector 78

3. Cost analysis 83

B. Qualitative 89

Chapter VI: Conclusions and Recommendations 114

References 118

Appendices

Appendix A: Articles and Papers

1. A Look at Some Assumptions in Special Education

2. Teacher Expectancy and Student Achievement

3. The Role of the Teacher Aide in Pupil Management

4. The Utility of Mental Age as a Predictor of Academic Achievement

5. Noise Levels in the Classroom

6. Light Levels in the Classroom

Appendix B: Test Manual and Scoring Materials for Behavior Analysis Scale - Educable (BASE)

Appendix C: Language Analysis

1. Analysis of Language Samples
2. Factors Affecting Language Indices
3. Changes in Language Indices as a Function of Race, Sex, and IQ

Appendix D: Testing Procedures and Copies of Forms Used in the Study

1. Analysis of Errors in Using ITPA
2. Modified WRAT Administration Instructions
3. Forms Used in the Study

Appendix E: Analytical Procedures

1. Automatic Interaction Detector (AID)
2. Some General Comments on the Relation of Multivariate Analysis to Repeated Measures Analysis of Variance

LIST OF TABLES

TABLES PAGE

1 Comparison of Average Gains by Multiple-Faculty, Large-Control, and Normal-Control Categories on General Achievement Tests, in Grade-Equivalent Years 13

2 Restrictions Imposed by Teachers in Becoming Part of the Project 31

3 Non-Linear Component of Time Trend for WRAT Scores 36

4 Analysis of Variance for WRAT Scores 39

5 Estimates of Effects 41

6 Correlations Among Average BASE Scores for the Three Test Points (Elements Above Diagonal Omitted) 47

7 Variance of BASE Scale Scores (Average of Three Test Points) 48

8 Multivariate Analysis of Variance for BASE Scales 48

9 Least Square Estimates of BASE Scale Means for the Time Effect and the Aide x Time Effect (Estimates for the Full Model) 49

10 Least Square Estimates of BASE Scale Means for the Class Size x Time Effect (Estimates for the Full Model) 51

11 Non-Linear Time Trend Analysis 52

12 Analysis of Variance of Film Indices 53

13 Estimates of Differences of Film Indices 55

14 Analysis of Variance for Absenteeism 59

15 Multivariate Analysis of Variations Among ITPA Sub-Tests 61

16 Analysis of Variance of Average ITPA Test Scores 62

17 Estimates of Effects for Average ITPA Scores 62

18 Scale Changes in Linguistic Analysis 63

TABLES PAGE

19 Grouping of 19 Revised Indices and Desirability of Increased Magnitude of Scores 66

20 Mean, Variances, and Standard Deviations of Composite Indices and Correlations Among Indices 67

21 Composition of Final Indices and Scale Modification on Each 67

22 Multivariate Analysis of Variance: Linguistic Indices 68

23 Least Square Estimates of Class Type x Size Differences and Standard Errors 69

24 Estimates of Time Trend Components 70

25 Least Square Estimates of Mean for the Class Type x Time Interaction 70

26 Least Square Estimates of Class Type x Time Parameters of Linguistic Indices 71

27 Estimates of Differences in "Linear" Component of Change Associated with Aides by Class Size Interaction 71

28 Estimates of Response by Index by Time in Presence or Absence of Aides and for Standard and Increased Size Classes 72

29 Estimates of the Linear Component of Time Trend for Acceptability by Cell 73

30 Raw Data for Cost Analysis Based on the Two Semesters in the 1967-68 School Year 84

31 Source Table for Analysis of Variance of Cost-Achievement Data 86

32 Two-Way Tables of Means and Frequencies of Cost-Achievement Data for the Three Factors of Class, Aides and Size. The Means are Based on Class Averages of the Dollar Cost of an EOB Pupil Achieving One Academic Year (10 Months) 88

LIST OF FIGURES

FIGURES PAGE

1 Reading Scores on the WRAT as a Function of Class for EMR Pupils Over Three Semesters 63

2 Spelling Scores on the WRAT as a Function of Class for EMR Pupils Over Three Semesters 64

3 Interaction of Arithmetic Scores on the WRAT as a Function of Class and Class Size for EMR Pupils Over Three Semesters 65

4 Reading, Spelling, and Arithmetic Scores on the WRAT as a Function of Aides for EMR Pupils Over Three Semesters 68

ACKNOWLEDGMENTS

The successful completion of this project was due to the efforts of a considerable number of people who appear in the listing below. Special acknowledgment must, however, be made to the project personnel, in particular Mrs. Sarah Scott, the on-site coordinator of the project. The cooperation of the Milwaukee Public School System in helping to expedite the project is appreciated.

List of People Who Were Involved in the Project

Co-Investigators

Blessing, Kenneth, Ph.D.
Blodgett, E. Donald, Ed.D.
Contrucci, Victor
Cook, John J., Ph.D.

Principals

Bolch, Howard H.
Feldman, Hollis L.
Fish, Herbert
Gordon, Grant
Guthrie, John
Graffenberger, Sara
Hahn, Robert A.
Halfast, Byron
James, Lloyd F.
Koettig, Margaret
Lipke, Melvin J.
Paul, Wayne E.
Powell, Earl F.
Rheinlaender, Ernest A.
Rose, Raymond W.
Tabla, Edward

Teachers

Allen, Joyce
Beard, Isreal
Bundy, Ida
Bright, Ellen
Buckner, Grace
Carriveau, Gene
Carter, Valeria
Clardy, Ralph
Cyrulik, Antoinette
Davis, Minnie
Drollinger, Barbara
Eichenbaum, Bertha
Uhlman, Patricia
Godwin, James

Project Personnel

Lind, Victor, Data Analyst
Miller, Karen, Project Secretar
Miller, Lillian, On-Site Secret
Scott, Sarah, On-Site Coordinator
Swanson, Mike, Cameraman
Sewell, Diane, Secretary

Schools

Alexander Mitchell
Byron Kilbourn
Walter Allen
Garfield Avenue
Robert M. LaFollette
Lincoln Avenue
Luther Burbank
Oliver Wendell Holmes
Carleton
Eugene Field
38th Street
Fourth Street
Victor L. Berger
William McKinley
James Whitcomb Riley
Henry L. Palmer

Teachers

Gray, Herbert
Gunn, Carrie
James, Dorothy
Klander, Suzanne
Maxwell, Mildred
Meriweather, Arvella
More, Virginia
Patton, Katharyn
Pittman, Daisy
Roberts, Douglas
Waldoch, Joyce
Williams, Anna
Wadyn, John

Teacher Aides

Brewer, Susan
Davenporte, Theresa
Edwards, Mary
Franklin, Velma
Hendrickson, Susan
Hinke, Martha
Holmes, Thelma
Johnson, Lourenia
Kleckly, Janie

Project Assistants

Anscher, William
Bange, Terry
Beeman, Rita
Bergren, Mary
Bernauer, Margaret
Blodgett, Judy
Boss, Lauretta
Bouma, Rita
Burns, Barbara
Christensen, Conrad
Crump, Richard
Crowder, Ellen
Daniels, Shelton
Dorn, Karen
Dettmann, Marilyn
Edwards, W. Richard
Falck, Glen
Feldman, Joan
Frank, Rochelle
Goldammer, John
Hauck, Sandra
Harper, Hoyt
Harrington, Dorothy
Hesse, John
Jones, Mrs. Frankie

Teacher Aides

Lieberman, Sharon
Miller, Bessie
Peters, Aurelia
Remus, Dorothy
Rusbuk, Mary
Stanley, Elnola
Swope, Lola
Taylor, Lena
Wagner, Judith

Project Assistants

Kaster, Roni
Kraft, Judy
Kunz, Deanna
Magner, Clem
Mandell, Judy
Martin, Lyn
McGruer, Joan
Mock, Carol
Neal, Harrison C.
Nowicki, Cammille
Osset, Dr. Martha
Pardo, Casilda
Pflaum, John
Plous, Helen
Pace, Rose
Raspotnik, Faye
Rosner, Martha
Sawyer, Dale
Shalit, Andrea
Sharski, Maureen
Starrett, Robert
Sweet, Gail
Townsend, Mary Ann
Triska, Julie
Van Wagnen, Margaret
Zimmer, Thomas

Consultants

Broskowski, Anthony, Ph.D., Clinical Psychologist
Brown, Betty, Language Specialist
Dever, Richard, Ph.D., Linguist in Mental Retardation
Fischbach, Thomas, Statistician
Heber, Rick, Ph.D., Psychologist
Jackson, John, Ph.D., Psychologist
Kitonga, Ellen Crowley, Ph.D., Linguist
Osmond, Louis, M.S. Supervisor
Olson, James, Ph.D., Special Education
Ramstad, Robert, Ph.D., Research Design
Scheerenberger, Richard, Ph.D., Administrator
Schwarz, Robert, Ph.D., Administrator
Scotton, Carol, Ph.D., Linguist
Wolf, Richard, Computer Programmer

The University of Wisconsin Computing Center provided additional support which was made possible through the National Science Foundation and Wisconsin Alumni Research Foundation.

SUMMARY

Research evidence regarding the use of teacher aides and varying class size in the education of the mentally retarded has been at best contradictory and at worst non-existent. The impact in terms of cost and possible benefits (or lack of them) of these aspects of the educational process warranted their close examination.

EMR classes were assigned, on a sequential random basis, to the cells of a 2x2x2x3 repeated measures design. The factors of the design were class (primary-intermediate), aides (present-absent), size (standard-standard + 30%), and time (three testing times spaced eight months apart). Dependent variables in the study were the following:

(1) Academic - WRAT subtests reading, spelling, and arithmetic.

(2) Language - Illinois Test of Psycholinguistic Abilities (ITPA)
- language measures derived from speech samples.

(3) Behavior - five indices of rated classroom behavior: disruptive behavior, attentive involvement, withdrawal, timidity, and succorance.
- three indices of classroom behavior derived from filming: noise level, activity level, and task-oriented behavior.

- absentee rate.

Analysis of the data was done using a general multivariate analysis of variance program and an automatic interaction detector analysis (AID). Cost-achievement and qualitative analyses were also done.

The following conclusions appeared warranted:

(1) The major impact of the aides alone (i.e., without regard to other conditions such as size of class, age level of children, etc.) appeared to be, first, a reduction in undesirable behavior in the classroom, and, second, an enhancement of the spelling achievement of the pupils.

(2) Class size alone had an effect on both arithmetic and classroom behavior. The standard (i.e., smaller) classes gained more in arithmetic and manifested less negative behavior than the increased classes.

(3) In general, what effects the factors of aides and class size had appeared to be most pronounced in the primary classes suggesting that administrative strategies other than the use of aides and manipulating class size should be developed in order to enhance the academic, behavioral, and linguistic functioning of intermediate EMR pupils.

(4) When questions of cost per unit of academic achievement were raised, the following recommendations could be made: either increased size without an aide, standard size without an aide, or increased size with an aide, (for all these conditions the average cost per pupil per ten months of academic achievement = $2,343) were to be preferred to the standard size class with an aide (average cost per pupil per ten months of achievement = $4,033).

(5) Qualitative analysis of interviews, conferences, and film material suggested that administrators should effect a fine balance between the primacy of either function or structure in the utilization of teacher aides. Too much structure, in terms of rigid role definition, teacher-aide, aide-parent, and aide-pupil relationships, etc., tended to nullify the full contribution that the aides could make. On the other hand, the aide's lack of familiarity in the educational setting dictated that some broad guidelines and training opportunities need to be provided.

Limitations of the study and its implications for administrative decision-making were briefly discussed.

Other immediate outputs from the project were:

(1) A behavior rating scale for the EMR (BASE).

(2) An approach to (and data from) language analysis in the retarded.

(3) Classroom films used for dissemination and teacher-training purposes.

(4) Analysis of:

(a) light and noise levels in the classroom
(b) teacher expectancy and achievement in the retarded
(c) certain assumptions in special education
(d) mental age as a predictor of achievement in the retarded

(5) A great amount of data, much of which is being subjected to various post priori analyses.

PRELIMINARY COMMENTS

At its inception the concerns of this study were relatively simple. Did having or not having an aide in the classroom, increasing or not increasing class size have a differential effect on the academic and language achievement, and classroom behavior of primary and intermediate educable mentally retarded children (EMR)? A number of factors entered, however, which resulted in an increase in the complexity of the study and an expenditure of effort on concerns not anticipated at the beginning of the project.

Specifically, assumption had been made regarding the availability and appropriateness of instrumentation to index the dependent variables. In several instances these assumptions were wrong, and instrumentation and techniques had to be developed during the course of the project. Thus, a behavior rating scale appropriate to the population of the study was developed. Stimulus materials and analytic techniques were developed to handle the problems encountered in indexing language change.

Additional concerns of considerable importance presented themselves after the project got underway. Since data relevant to these concerns could be collected with a minimum of expense and effort, the boundaries of the study were extended. Among these concerns were teacher expectancy as it relates to achievement, meteorological and other environmental factors and their influence on behavior of pupils, and the validity of certain assumptions in special education.

Monitoring of the project by means of film was done. The films provided a number of dependent measures as well as the footage necessary to make a 23-minute film which told the story of the project. This film is available for general use with other parts of the footage being available in the local Instructional Materials Center for professional instructional use. Here again the boundaries of the project were extended to encompass what were felt to be worthwhile and relevant undertakings.

Most of the additional analyses, test-taking instructions, etc., are referred to in the main body of the report with the actual write-ups being in the Appendices in article or test manual format.

Very little of a theoretical nature is included in the report as the applied thrust of the research precluded too much emphasis on the theoretical. This could be seen as a deficiency, but it is hoped that the findings are of sufficient validity to be incorporated into theoretical frameworks.

CHAPTER I

PROBLEMS UNDER CONSIDERATION

A. Teacher aides and class size as factors in pupil achievement and classroom behavior.

The increasing numbers of mentally retarded children for whom additional teaching personnel must be provided within the public schools has created a need for emergency steps to be taken leading toward adequate educational arrangements for these pupils. New innovative approaches are required to resolve this problem of serving the large numbers of mentally retarded pupils being identified as potential special class enrollees at the present time.

At present, a substantial portion of the professional special educator's teaching day is spent with duties which do not require professional preparation. A large number of these duties could be carried out by individuals without experience or training in professional education. With careful selection and specific orientation, lay persons could conceivably be used as adjunctive teacher aides, thus permitting the teacher of the retarded to serve more children and to spend a greater portion of her time in activities directly related to her knowledge, skills, and highly specialized training.

A related issue to the continuing problem of educating mentally retarded children has been the impact of traditionalism in administrative grouping practices and levels of required competency prescribed for personnel working in these programs. In the past by some empiric method, classroom enrollments for the retarded were restricted to approximately half the size of a regular classroom. Acceptance, generally, of the need for certain professional competencies and restrictive class enrollment sizes has led to the utilization of the following two criteria for establishing and operating a unit:

(1) a teacher trained and certified in the education of the mentally retarded; and

(2) a restrictive limitation on enrollment based on the premise that these children exhibit lower learning rates, greater behavioral and learning problems thereby requiring greater individualized attention.

The purpose of this study was to determine the effects of class size and the use of teacher aides on pupil behavior and teacher behavior in EMR classes at primary and intermediate grade levels. Effects on pupil behavior were determined by contrasting the educational and linguistic achievement and social behaviors in the several instructional arrangements.

The study was designed to demonstrate the feasibility of increasing traditional primary and intermediate enrollments without impairing, and if possible, actually improving the learning efficiency and pupil behavior of the specific classroom units incorporating teacher assistants as adjuncts to the instructional program.

B. A behavior rating scale for the retarded

The need for a behavior rating scale standardized on a population of retarded pupils became evident as the project got underway. One of the dependent measures which was expected to be influenced by the experimental manipulations was that of classroom behavior. Accordingly, such a scale was developed during the course of the project. Appendix B contains the test manual and scoring procedures for this scale which has been named Behavior Analysis Scale - Educable (BASE).

C. An approach to language analysis in the retarded

Differential growth in language was another dependent variable which was expected to be affected through varying use of aides and class size. The Illinois Test of Psycholinguistic Abilities (ITPA) (Kirk and McCarthy, 1961) was initially considered appropriate for this purpose. However, factors such as cost of administration, difficulty in getting tester agreement on scoring responses from the target population and the fact that pupils representative of the study population were not included in the test's standardization sample suggested significant weaknesses existed in the use of this test in this study. While the ITPA was used in the study, the number of pupils tested was restricted (a 30% random sample) as was the number of testings (two versus three for the other measures).

There was arising at the inception of the project a considerable demand for simpler evaluation techniques in those federally funded demonstration programs concerned with improving the language development of the retarded. This consideration, plus the limitations apparent in the ITPA, suggested that the development of easily obtainable indices of language development could have considerable value, not only to this study but the field as a whole. The results of such an undertaking along with subsidiary analysis of the indices are contained in Appendix C.

D. Filming in the classroom

A frequent criticism of research is that little attempt is made to widely disseminate the findings, particularly in a media and a format readily understood by the on-line administrator or supervisor. A partial solution to this problem in the present study was to do representative filming throughout the length of the project. A film entitled "Teacher Aides" based on this footage is on deposit at the U.S.O.E. Loan copies are also available from the University of Wisconsin's Bureau of Audio-Visual Instruction in Madison.

The fairly recent development of compact and quality cameras and video-tape apparati has resulted in their wide-spread use in the field of education. At least two broad forms of use (excluding dissemination) of the large amount of footage can be detected. One is the purely educational role in which the learning teacher views her own performance and thereby gathers insights which result in improved modifications of her teaching techniques; or she views specified sequences of others at work illustrating a better way of doing things or perhaps as a point of departure for discussion. The present study used the films for this purpose by having them available to the teachers and aides at all times (they were viewed infrequently) and by using them in the inservice sessions (they had no choice!). At the present time, the film segments along with detailed descriptions of content and quality are housed in the Special Education Instructional Materials Center on the University of Wisconsin's Madison campus. The films are being used by special education personnel for pre- and inservice training of teachers and aides.

E. Influence of meteorological factors on behavior in the classroom

A frequently made observation of educators is that meteorological factors such as barometric pressure, wind speed and temperature have a marked effect on pupils' behavior in the classroom. That this could be the case is not surprising. However, the potency of the meteorological effect on a particular measure of behavior can be questioned. For, stronger causes such as an empty stomach, lack of sleep, or even drugs could also be hypothesized which would wipe out what would seem to be the more subtle effect of the meteorological factors. The nature of the present study allowed us to take a cursory look at this phenomenon. At the same time the filming segments were being picked up, readings on barometric pressure, wind speed, temperature in and outside were obtained. These were then used to predict the behavioral indices obtained from the concomitant film segments. The results of this analysis have not been written up at this time. However, papers on other environmental factors, noise levels and light levels are in Appendix A.

F. Teacher expectancy and achievement in the mentally retarded

At the start of this project (September, 1967), teacher expectancy was an interesting research finding. By the end of the experimental part of the project (June, 1969), it appeared to be an accepted fact. At this time (August, 1970), the manner in which the phenomenon has been substantiated is receiving rather sharp criticism (Thorndike, 1968, 1969). Specifically, the appropriateness of the scale of measurement for the dependent variable IQ scores is questioned. Thorndike also makes the point that if the expectancy phenomenon does exist, its effect would be most apparent in those areas which are directly teacher based and school dependent.

As though in anticipation of Thorndike's concerns, at the beginning of the present project teacher expectancy for their pupils' achievement in three almost solely school-dependent areas was obtained. These were academics (subject matter skills), social behavior (class adjustment), and language (self-expression, written and verbal). An examination of the relation between teacher expectancy and gains in these three areas over three semesters is contained in *Appendix A*.

G. A look at some assumptions in special education

There are a number of assumptions made by special educators which have been subject to little, if any, examination. These assumptions influence decisions about program development for handicapped children and, as such, need to be examined. These are:

(1) It is assumed that there is a regression or loss of academic skills over the summer.

(2) It is assumed that summer school programs have a positive effect on academic learnings.

(3) It is assumed that changing a child's teacher during the school year has an adverse effect on the learning of academic subjects.

(4) It is assumed that the more a child is absent from school, the less he will learn of academic subjects.

A continuing search of the literature has failed to reveal any research pertaining to these assumptions as they related to the EMR. This lack would seem to suggest one of three things: (1) the assumptions are extremely self-evident, (2) they are continually supported by the observations of the personnel involved, or (3) it is meaningless to undertake research studies in the four areas because of the host of other variables which influence learning in children.

Data from the present project were used to examine the four assumptions. Appendix A contains a reprint of an article encorporating this analysis.

H. The utility of MA as a predictor of achievement

The relationship between CA, IQ, and expected achievement based on MA in the retarded was examined. This endeavor has been accepted as an article for publication. (See copy of article in *Appendix A*.)

I. Cost analysis

Cost analyses were also conducted to determine the economic as well as the instructional efficiency of the teacher-aide plan, particularly in the experimental units with increased enrollments. Comparisons of traditional unit costs vs. the teacher-aide unit costs were made to examine the fiscal economy of this latter approach.

2. Other considerations

This section deals with problems that have been considered yet not developed to their fullest for any one of several reasons, viz., cursory inspection of the data suggested nothing of value or other work pressures prevented the authors from spending the time necessary to produce a fairly respectable finished product within the deadline set for this final report. Some of these problems are as follows along with their disposition:

(1) The Survey of Interpersonal Values (Gordon, L.V., 1960) was given to teachers and aides to determine if an analysis of the profiles could allow one to use the instrument to help to predict which teacher and aide dyads could be expected to work effectively together. An examination of the data was unfruitful mainly because of the small N involved and the difficulty in getting objective agreements as to who the effective and ineffective dyads really were and why.

(2) The relationship of the various language indices (see Appendix C) as predictors of IQ was analyzed. The write-up of this analysis has not been completed at this time.

CHAPTER II

REVIEW OF THE LITERATURE

As the previous chapter suggests, the number of problem areas considered during the project was considerable. However, this review of the literature is restricted to those studies which have direct bearing on the effects of aide utilization and class size. Reviews of the literature for the other areas are incorporated in the individual write-ups contained in the appendices.

The review of the literature takes into consideration both research and non-research reports relevant to: (1) a general overview of the teacher-aide program; (2) opinion surveys concerning teacher aides; (3) effectiveness of teacher aides as measured by demonstrated pupil progress; and (4) effect of classroom size (i.e., pupil enrollment) on student progress. Each area will be considered independently.

Overview of the Teacher-Aide Program

The purpose of this section is to provide an overview of the concepts underlying the teacher-aide program in an attempt to clarify, or at least delineate, some of the variables essential to research.

According to available reports, purposes of a teacher-aide program can be collated under two relatively distinct categories, each of which have administrative and research significance. These include: (1) purposes which are primarily teacher-pupil oriented, and (2) purposes which emphasize needs of a given subpopulation from which teacher aides are recruited and trained.

Teacher-pupil oriented purposes

Most statements concerning teacher-aide programs identify the role of the teacher aide as one which enables the regular classroom teacher to devote more time to professional activities. Four primary goals are consistently identified (e.g., Shipp, 1967; Emmerling and Chavis, 1966; California, 1968; Editorial, 1967; Riessman and Gartner, 1969):

(1) To enable the teacher to devote more of her time to her teaching responsibilities;

(2) To enable more individualized instruction;

(3) To relieve the teacher of clerical duties; and

(4) To enrich child-adult relations, i.e., a second adult gives the child an alternative person to relate to, to work with, and to serve as a model.

These goals, originally developed for "normal" and Headstart classroom situations, also serve as a basis for utilizing teacher aides in special classrooms (e.g., Blessing, 1967; Cruickshank and Haring, 1957; and Cross, 1968). Blessing (1967, p. 157), for example, emphasized the notion that ". . . an aide in a classroom can relieve the teacher of a tremendous amount of work and facilitate the professional task of the instructor."

An additional purpose for introducing a teacher-aide program into special education has been noted - to increase the number of mentally retarded students per teacher. Cross (1968, p. 75) stated that with the advent of the aide, "class enrollment was raised (from ten) to fifteen students." Blessing (1967) also makes reference to this potential value of teacher aides. It is interesting to note that this concept, which was evidenced only in literature concerning special education, seems in substantial contradiction to the position advanced by the National Educational Association: "The use of a staff approach to teaching should decrease the student-teacher ratio, not increase it" (Edelfelt, 1967, p. 13).

Purposes related to meeting needs of a given subpopulation

Several systems report that the primary function of their teacher-aide program was to provide supplemental financial support to ADC mothers and retired persons (Phipps, 1967); to prevent delinquency (MacLennan, 1966); and to provide incentive for teenagers to remain in school (Levy and Crosser, 1967). Goodman and Arnold (1967) reported a successful effort to recruit and train 43 unemployable persons, aged 18 to 53, to serve as assistants and aides in a number of programs for the mentally retarded, including day care centers.

While the intent of such programs ultimately is to serve children within an educational environment, the primary goal, and subsequently any evaluation of success or failure, rests with the progress of the aide rather than with the child.

Duties and qualifications

Beyond the basic premise that the primary function of the teacher aide is to enable the teacher to devote more time to teaching duties, there is little common agreement. In fact, "teacher aide" appears to be a generic label applied to any person other than the teacher employed in the classroom.

Duties vary considerably, ranging from simple clerical tasks to tutoring students in highly technical areas of knowledge. Primary categories of duties, however, as defined by Croft Educational Services (1968) include: (1) clerical, (2) housekeeping, (3) non-instructional, (4) audio-visual assistance, and (5) semi-instruction. The latter category includes such tasks as teaching good manners, reading spelling and vocabulary lists, assisting teachers in presenting special units, checking pupils' seat work, and reading and story telling.

Results of a study conducted by the National Educational Association (1967) indicated that the ten most widespread duties of paid teacher aides, arranged in descending order of frequency, were (1) duplicating tests and other materials; (2) helping with classroom housekeeping; (3) typing class materials, tests, etc.; (4) setting up audio-visual equipment and other instructional materials; (5) helping with children's clothing; (6) supervising the playground; (7) correcting tests, homework, workbooks, etc.; (8) reading aloud and storytelling; (9) assisting in the school library; and (10) collecting money from pupils. In essence, these duties are primarily supportive and non-technical in nature. Such duties have been reported by authors representing the use of teacher aides in elementary and secondary school systems (Emmerling and Chavis, 1966; Nelson, 1967; and Shipp, 1967) as well as special education (Blessing, 1967, and Cruse, 1966).

Other school systems, however, have deviated from the normal roster of duties in which the teacher aide plays a secondary role with regard to direct student instruction. Shields (1966) used teacher aides (undergraduate students) in a student teacher capacity. In other words, teacher aides assumed a direct teaching responsibility for subjects in which they were qualified. Levenson (1967) reports that, in California, aides were employed to tutor foreign-speaking children in English. Thomas (1967) reports that the teacher aide's primary responsibility was to assist in the oral language development of young children. In some schools, teacher aides provide counseling to parents. In such instances, teacher aides were usually mothers recruited from the social-economic area served by the school (Goldstein, 1966; and Teacher stretchers, 1967). It is quite evident that teacher aides, in some instances, are assuming a direct teaching role, and they are assuming duties normally considered within the purview of the school social worker.

With regard to the distribution of teacher aide time, a NEA editorial (1967) indicated that 30 percent of the teacher aide's time was devoted to routine clerical duties; 26 percent to supervising and monitoring activities; 22 percent to providing personal attention to students; and 4 percent to meeting with and/or counseling parents. No background data were provided. Range in number of hours served was one and one-half per day to a full week schedule.

Persons employed as teacher aides vary considerably in background. Requirements range from "18 to 21 years of age, from poor families living in high delinquency areas, free of infectious diseases, and not subject to pending court action" (MacLennan, 1966) to college graduates. Included are teenagers (e.g., Lamp and Crossman, 1967), mothers (e.g., Shipp, 1967 and Goldstein, 1966) retired persons (e.g., Shipp, 1967), former Peace Corpsmen (e.g., Bunn, 1967), and college students (e.g., Lamp and Crossman, 1967).

Administrative-research Implications

Administrative acceptance of the teacher-aide program should be based on realistic, sound research. The cost of teacher-aide program on a national basis represents a multi-million dollar effort. To illustrate, according to the NEA survey conducted during the 1965-66 school year, paid teacher aides contributed 500,000 hours per week! At an average reimbursement rate of $2.00 per hour, this represents a total expenditure of 32 million dollars per year. The total number of paid aides was approximately 30,000 (NEA, 1967). Today, there are 200,000 aides and by 1977 Riessman and Gartner (1969) estimate that the number of aides will reach 1.5 million.

Continued support for this program must be predicted upon unequivocal evidence that the contributions of the teacher aide are of significant value. If the primary purpose of the program is to permit the classroom teacher to devote her full-time energies to professional instruction, then the only *realistic evidence of significance is accelerated student growth.* A secondary justification might involve an increased student-teacher ratio in the event that larger classes, with the assistance of a teacher aide, do not result in depressed pupil achievement.

If, as stated previously, the primary goal of the teacher-aide program is to provide employment-training opportunities for a "marginal" subpopulation, then criteria of success must relate to the socio-personal and/or employment adequacy of individuals recruited. Student progress, in this instance, is of secondary importance.

Inasmuch as the teacher-aide program is a rather nebulous operational phenomenon, research projects should consider at least the following variables with regard to the aide: (1) specific tasks to be performed, interpreted in terms of measurable outcomes; (2) education level(s); and (3) number of serviceable hours. More sophisticated efforts should include reference to aides' sex, age, personality, and socio-cultural background in relation to the student population being served.

Evaluation of student progress should include at least a series of reasonably spaced time intervals and relevant objective tests.

Opinion Surveys

Much of the support for the teacher-aide program as reported in the literature is based on teacher opinion surveys. In each case, including the special class, the teacher aide has been received with enthusiasm by the regular classroom teacher.

To illustrate, the NEA (1967 a, p. 37) reports that "an overwhelming majority (nine in ten teachers) indicated that having teacher aides was helpful and more than half said that it was a *great* assistance. Teachers in small school systems were the most enthusiastic about the assistance provided by their aides." Distribution of responses is shown below:

Response category	Percent teacher response
Great assistance	51.4%
Some assistance	38.4%
Little assistance	9.5%
No assistance	0.7%

In a study to compare relative effectiveness of six methods of teaching beginning reading to approximately 4,000 students in 19 inner city elementary schools, Rankin (1967) included teacher evaluation of teacher aides. Of the 111 of 112 teachers who completed the questionnaire:

69 teachers believed the aides to be of "much" value in helping the child to understand and carry out assignments.

53 teachers believed the aides to be of "much" value in teaching the child a specific skill or concept.

66 teachers believed the aides to be of "much" value in listening to the child read.

71 teachers believed the aides to be of "much" value in checking a child's work.

67 teachers believed the aides to be of "much" value in encouraging a child to improve his behavior or effort.

80 teachers believed the aides to be of "much" value in helping a child practice a skill.

Rankin (1967, p. 26-27) concluded that, "according to the teachers' reports, the lay aides provided a wide variety of useful services in support of the instructional program.... Some of the services involved direct assistance in teaching individual children and in guiding group learning experiences. Other services included general supervision of children, preparation of instructional materials, and the performance of various clerical and record-keeping tasks."

Also included in the study were the results of time sample observations by the researchers which indicated that non-project teachers (without aides) devoted 68 percent of their time to instructional matters as compared to 75 percent of the experimental teachers, (a difference of dubious practical significance).

With regard to special education, Green (1967) reported that among Chicago teachers of trainable children with teaching aides, 95 percent indicated that they "liked" having the aide. No specific data concerning method, sample, or questions asked were included.

Blessing (1967, p. 109) in his review of the Cruickshank and Haring study which utilized teacher aides in three types of special classrooms (blind, gifted, and mentally retarded) reported that the results:

> . . . while highly subjective, generally favored the use of teacher assistants in the special education classes. All of the 19 teachers who had total or shared responsibility for a teacher assistant reported that they were able to do more creative planning, provide more materials, give more individual attention to the children, and, in general, increase the quality of their instruction. The teachers reported that the parents unanimously felt their children had gained from the Teacher

Assistant Plan. Also, all of the administrators of the school involved were receptive to the program and felt that their teachers had done a better job of teaching as a result of having had an assistant.

Though opinion surveys are subject to a host of difficulties, primarily involving reliability, there would appear to be sufficient data to indicate that the teacher-aide program is well received by educators. There is some evidence to indicate that, in fact, the teacher aide does reduce rather significantly the amount of time normally devoted by the classroom teacher to clerical duties. According to the results of a five-year study conducted in 23 Michigan schools (Behavioral Science Notes, 1967), aides cut teachers' time on paper correction by 89 percent, classroom monitoring by 83 percent, and were of considerable assistance in the preparation of various reports.

The major question is not one of validity of results within the realm of research design, but rather the administrative priority one wishes to assign to such data. In other words, do the expressed opinions of teachers *per se* constitute sufficient justification to initiate a large-scale, relatively expensive teacher-aide program?

Effectiveness of Teacher Aides as Measured by Demonstrated Pupil Progress

In contrast to the large number of articles prepared concerning the philosophical and programmatic aspects of the teacher-aide program, few concrete research projects have been conducted. None was located which dealt specifically with the mentally retarded.

One study reported by the California Office of Compensatory Education (1968), concerned the influence of teacher aides in a reading program intended to serve migrant children. Progress of 24 resident children (i.e., children enrolled in a regular school program) were compared with 24 migrant children who received extra reading instruction from two teacher aides for a period of four months. Both groups were pre- (November, 1967) and post-tested (March, 1968) on the reading subtest of the Wide Range Achievement Test (WRAT) and the Peabody Picture Vocabulary Test (PPVT). Post-test analysis of *PPVT* data revealed that the raw scores of the 24 migrant children improved by 5.4 points, while raw scores for the 24 resident children only attained a mean gain of 4.6. On the *WRAT*, scores of the migrant children showed a mean gain of eight months compared with the resident students' gain of only four months. It was concluded (p. 13) that "the improvement and test scores of the migrant students were greater on both tests than that of the residents. This difference was attributed by the evaluators to the special help provided to migrant students." Specific characteristics of both migrant and resident children were not presented. Absence of a migrant control group renders the conclusions somewhat spurious. Also, the teacher aide was active in a classroom for only a very short period of time.

Other studies, however, have demonstrated that the teacher aide may play a significant role in assisting the young child in the development of reading and number readiness. To illustrate, based on the performance of 234 Minneapolis school children, using the Metropolitan Reading Readiness Test, pupil learning was 50 percent greater in kindergarten classes with a paraprofessional than where there was no aide (Riessman and Gartner, 1966).

Howell (1939) reported the results of the Yale-Fairfield study of elementary teaching which included both teacher aides and varying class sizes. Thirty-six groups (or classes) of third and fourth grade students were distributed equally across three class categories: 12 groups, multiple-faculty (with aides); 12 groups, large control; and 12 groups, normal control. Large control groups consisted of classes with a mean class size of 29.4 as compared with normal control groups with a mean class size of 24.8. Mean class size for the multiple-faculty was 30.4 students.

Results in academic achievement as measured by the Metropolitan Achievement Tests and the California Achievement Tests are shown in Table 1. Gains were measured on the basis of pre- and post-testing with an interval of one year (i.e., May 1936 - May 1937). Analysis of variance between the multiple-faculty and large-control and between large-control and normal-control categories yielded no statistically significant differences. Thus, it was concluded that progress was not significantly improved with the appearance of an aide, nor did class size make any appreciable difference. It should be noted, however, that the range between the three groups, based on average class size, was only 5.6 students, i.e., 24.8 to 30.4.

Table 1

Comparison of Average Gains by Multiple-faculty, Large-control, and Normal-control Categories on General Achievement Tests, in Grade-equivalent Years (Howell, 1939, p. 29)

	Category		
	Multiple Faculty	Large Control	Normal Control
Reading	1.14	1.08	.88
Arithmetic	1.13	1.06	.90
Spelling (i)	.96	.92	.89
English and grammar (ii)	1.40	1.04	1.07
"Total achievement"	1.13	1.05	.94

(i) Grades 3 and 4 only: ten groups in each category.
(ii) Grade 4 only: four groups in each category.

Anastasiow (1964) reported an interesting study in which a second teacher (not a teacher aide) was assigned to each of six kindergarten groups serving primarily former children of Project Headstart. The overall results of the study in which the supplemental teachers appeared in the classroom for a three-month period made gains on the three tests employed than would have been expected by chance alone: 92 of 116 children showed gains in object identification; 81 of 132 children made gains as measured by the Draw-A-Man Test, and 123 of 176 children made gains in language as measured by a locally prepared Language and Speech Evaluation Test. Again, there were no control groups--nor was there any specific indication as to quantitative magnitude of the gains and their practical significance, if any.

These few studies represent the extent of the literature located concerned with assessing effectiveness of teacher aides with specific regard to student progress. Conclusions concerning the efficacy of teacher aides as measured by student progress cannot be made on the basis of the presented studies. They are too few, and they possess major limitations, especially with regard to the adequacy of control, sample size, and the identification of teacher aides by qualifications or responsibilities. The single conclusion that can be advanced, however, is that considering the number of teacher-aide programs which have been established plus the considerable investment in such projects with regard to teacher time, in-service training, and financial obligations, the paucity of well-conceived, effectively implemented research studies is remarkable.

Effect on Classroom Size (Pupil Enrollment) on Student Progress

With regard to the relative size of class and pupil achievement, a recent publication by the Research Division (1968) of the National Education Association presents a rather thorough integration of 81 studies in this area. Though a large number of studies have dealt with such aspects as teacher attitudes on class size, the variety of activities offered in small versus large classes, and individualized instruction in small groups, only one study was concerned specifically with achievement level differences among elementary school children enrolled in small as compared with large classes. This study, which examined the effect of class size on the reading achievement of first grade pupils, was reported by Frymier in 1961. Subjects were pupils on whom complete achievement data were available in six large (more than 36 pupils) and nine small (fewer than 30 pupils) first grade classes. A total of 203 pupils in large classes and 219 pupils in small classes were studied. It was reported that, for all practical purposes, the two groups were similar in terms of visual and auditory acuity and physical health. Pupils in small classes were slightly older than pupils in large classes (mean age of 81.0 months compared with 79.1 months). Readiness tests administered at the beginning of first grade showed pupils in large classes to be significantly above pupils in small classes at the one percent level of confidence. There were slight, "but probably significant" differences between the two groups of teachers. All teachers had B.S. degrees, one teacher of a small class had a M.S. degree. Teachers of large classes appeared to be slightly more experienced.

Achievement was measured by (1) reading achievement test scores at the end of the year, (2) mean grade placement, and (3) retention and promotion rates. On the *Williams Primary Reading Achievement Test*, large-class pupils had a mean raw score of 19.21, and small-class pupils had a mean raw score of 22.58, a difference significant at the .001 level of confidence. Pupils in large classes had a mean grade placement of 1.62 at the end of the first year, compared with 1.75 years for small-class pupils. Differences between retention rates for large classes (23 percent were retained) and small classes (17 percent were retained) were not significant. Frymeier concluded that there was clear evidence that class size influenced achievement in reading for these first grade students.

Howell's study, previously reported, showed no statistically significant differences as a result of class size discrepancies. However, as noted, size differences were quite small.

Class size, with regard to the education of educable mentally retarded, has, for many years, been between 10 and 15 students per class per teacher. Though some variances are observed as a function of age, rarely has the class size exceeded 15. As indicated by Kirk and Johnson (1951, p. 126), in their classic text concerning the education of the retarded child:

> *The younger the children the smaller the class.* A preschool class should not have more than a ratio of five or six children to one teacher, whereas a class of adolescent children in the secondary school, where many of them are sent to other classes, could probably have a ratio of fifteen to twenty children per teacher. State regulations which set a lower and upper limit should differentiate between the younger and the older children.
>
> *The more homogeneous the class the larger it can be, and vice versa, the more heterogeneous the class, the smaller it should be.* When a heterogeneous class of children from eight to sixteen years of age is organized into one special class for the mentally retarded, the class should be relatively small. It is more difficult for a teacher to organize for such a class than for a class of children ranging in age from eight to ten. Class size, therefore, should be determined by heterogeneity and age, rather than be a set standard.

There is no research evidence pertaining to the accuracy of these observations. Size of special education classes appears to have resulted from a combination of teacher-administrator compromise, embellished with psychological justification and tradition. There is no study which either supports or does not support any hypothesis concerning relative class size as a function of mental level, level of intelligence, level of adaptive behavior, chronological age, or heterogeneity. The first special education class was established in the United States in 1878 (Kanner, 1964).

Conclusions

Only one conclusion can be reached concerning research in relation to either teacher-aide programs (for normal as well as exceptional child) or the effect of classroom size on student achievement. Educators and educational administrators, in spite of recent efforts to train a cadre of research personnel, have failed consistently to examine systematically and comprehensively the effectiveness of programs for which they are responsible. Evidence supporting teacher-aide programs, especially in regard to student progress, is, for all practical purposes, non-existent. Therefore, it is essential that the teacher-aide program be studied in depth with regard to its influence on student achievement as well as its potential value in terms of increasing student-teacher ratios.

CHAPTER III

DESIGN AND ANALYSIS CONSIDERATIONS

Design

The major objective of this study was, through experimental manipulation, to ascertain the differential effect of the use of teacher aides and two levels of class size on the academic achievement, language growth and classroom behavior of primary and intermediate class educable retarded. Possible interactions among these three factors were also sought. Cost analyses were also conducted to determine the economic as well as the instructional efficiency of the teacher-aide plan, particularly in the experimental units with increased enrollments. Comparisons of traditional unit costs vs. the teacher-aide unit costs were made to examine the fiscal economy of this latter approach.

The above objectives can be considered the quantitative aspect of the study.

Outcomes of a qualitative nature were also considered. Here the subjective reactions of teachers, aides, principals, and administrators were analyzed.

The basic design of the study was a 2x2x2x3 repeated measures design. The layout of the design was as follows:

Factor
Name

Class	Intermediate				Primary			
Aides	No Aides		Aides		No Aides		Aides	
Size	Inc.	Stan.	Inc.	Stan.	Inc.	Stan.	Inc.	Stan.
Time T_1	18.0	14.5	21.3	13.3	18.0	9.7	18.0	11.5
T_2	21.0	15.5	21.0	13.5	17.0	11.3	17.3	11.5
T_3	19.5	14.0	20.0	12.5	16.0	10.7	15.0	11.5

As the design shows, the four factors were *class*, a status factor at two levels - intermediate (10-13 year olds) and primary (7-10 year olds); the two manipulated factors, *aides* at two levels - with and without an aide; and *size* at two levels - standard size and standard size plus 50%; and the one temporal or repeated measures factor at three levels spaced eight months apart.

The numbers in the body of the design show the average number of pupils per class in each of the cells and illustrates the extent to which the size factor was controlled over the course of the study. Collapsing across the aide and time factor provides a more condensed illustration:

Intermediate		Primary	
Increased (Standard + 30%)	Standard	Increased (Standard + 30%)	Standard
20.0	13.9	16.9	10.9

On the average the target size of half again as many children for the increased class size was maintained quite well. An increase of 30% for the intermediate classes from the standard of 14 would be 21, and we managed an average of 20. An increase of 30% for the primary class from a standard of 11 would be 16 or 17, and we attained 17.

Dependent Variables

In their broadest sense the dependent variables in the study were three - academic achievement, classroom behavior, and language development. The problem of how to index or measure these areas loomed rather large for our particular population and in light of the resource and time constraints under which the project had to operate. Our solution was to view the whole process within the context of the geological model suggested by Webb, et al in their book Unobtrusive Measures: Nonreactive Research in the Social Sciences:

> "The 'outcropping' model from geology may be used more generally. Any given theory has innumerable implications and makes innumerable predictions which are unaccessible to available measures at any given time. The testing of the theory can only be done at the available outcroppings, those points where theoretical predictions and available instrumentation meet. Any one such outcropping is equivocal, and all types available should be checked. The more remote or independent such checks, the more confirmatory their agreement."

That is, the measures we obtained were viewed as samples of "outcropping" of the underlying substrata of academic achievement, language functioning, or classroom behavior.

The "substratum" of academic achievement was indexed by the three subtests - reading, spelling, and arithmetic of the Wide Range Achievement Test (WRAT) (Jastak and Jastak, 1965). For our purposes, this test had both positive and negative attributes. On the negative side, it was insensitive to small changes and probably for this reason not too reliable when most of our pupils, by answering one question correct, raised their achievement one month. On the positive side, it was economical to administer and it did extend far enough down to pick up a measurable score on all the pupils in the project.

The "substratum" of language development was indexed in two ways. The Illinois Test of Psycholinguistic Abilities (ITPA) (Kirk and McCarthy, 1961) with its nine subtests and total score was used with a smaller random sample of pupils in the project. The considerations in using this test perhaps suggest the utility of the ecological model. This test was not standardized on children representative of our sample in terms of either age or race. Thus, many responses could not be handled by the manual so our own scoring criteria were set up and used consistently with the expected [illegible]. Only raw scores were used in the analysis. Under these circumstances we see the scores as not a one-to-one measure of anything, but as an index of a substrata-language development.

A second index of language development was obtained over the course of the project by recording three-minute language samples of the pupils as they told stories about selected Charles Addams' drawings. These samples were then analyzed and 29 indices derived. This procedure along with supplementary data analyses is contained in Appendix C.

The "substratum" of classroom behavior was also indexed in three separate procedures. First, a behavior rating scale was developed during the course of the project. This rating scale, Behavior Analysis Scale - Education (BASE), consists of the five subtests, disruptive behavior, attention involvement, stability, acceptance, withdrawal.

The second aspect of behavior indexed was that displayed during filming sequences which went on during the project. These indices tapped total classroom behavior as opposed to the individual child's behavior. They were: (1) activity level in the classroom, (2) non-task oriented behavior, and (3) noise level in the classroom.

The third index, absentee rate, was obtained from the school district files.

Analyses

The strategy in the analyses of the data fell into two approaches. The first was to test the null hypotheses that there were no significant differences between the factor effects and among their interactions with time in terms of the components of the dependent variables conceptualized as "substrata." Specifically, the null hypotheses was that there were no significant effects due to the main effects and their interactions across time on the following groupings of dependent variables:

(1) the WRAT subtests reading, spelling and arithmetic indexing academic achievement;

(2) the ITPA subtests indexing language developments;

(3) the 29 language measures indexing language development;

(4) the five BASE subscales indexing classroom behavior;

(5) the three film measures indexing classroom behavior;

(6) the absentee rate indexing classroom behavior.

The statistical technique of choice with these hypotheses was that of multivariate analysis of variance given fairly widespread application by Bock (1963); Bock and Haggard (1968); Finn (1969). The multivariance computer programs developed by Finn (1968) allow a detailed analysis of the significance of the subtests which make up the multivariate analysis. Thus, if the multivariate analyses were significant, the individual subtests were examined to determine which one contributed most to the overall significant effects.

The second approach in analysis was explored because some of the assumptions of the first approach were violated to a certain extent. The problems encountered in sampling, etc., are described in the next chapter. The procedure used in the second phase of analysis is called AID (Automatic Interaction Detector) and is on file with the University of Wisconsin Social Data and Program Library Service. Linearity and additivity assumptions are not required. The model asks the question:

> "'Given the units of analysis under consideration, what single predictor variable will give us a maximum improvement in our ability to predict values of the dependent variable?' This question, embedded in an iterative scheme is the basis for the algorithm used in this program. The program divides the sample, through a series of binary splits, into a mutually exclusive series of subgroups. They are chosen so that at each step in the procedure, their means account for more of the total sum of squares (reduce the predictive error) than the means of any other equal number of subgroups." (Sonquist, 1970)

Using this model and program allowed us to ascertain in a more complete sense what factors were predictive of the academic achievement of the pupils over the three semesters of the project. The dependent variables in this analysis were the reading, spelling, and arithmetic gain scores. The predictor variables were the three design variables of class, aides, aide plus teacher, IQ, SES, teacher sex, race, and pupils' sex.

Further discussion of AID and the multivariate analysis of variance as it relates to repeated measures design is contained in Appendix E.

CHAPTER IV

METHOD

A. Selection and Characterization of the Sample

Teacher Selection

Available information at the beginning of the project suggested that a pool of some 80 teachers could be recruited into the project out of the 128 or so EMR teachers in the city of Milwaukee. Twenty-four from this pool were then to be randomly assigned to the eight cells of the design, i.e., three to a cell. Fall inservice meetings were attended by the project personnel, and the project was explained. In spite of a $150 per semester enticement and the likelihood of having an additional person to help out in the class, only 23 teachers opted to get involved. Moreover, over half of those who indicated willingness to participate would only do so subject to certain restrictions such as:

(1) Would not take an aide under any circumstances, but would take a chance on having a larger class size,

(2) Would take a larger class but only with an aide,

(3) Would come into project but only if she could go with her present class.

Table 2 displays the restrictions imposed by the various teachers. It is apparent that a completely random assignment to the cells within the primary and intermediate teacher groups was not possible. In an attempt to make the allocation of the teachers as unbiased as possible, a sequential random assignment was done. That is, those teachers with the most stringent restrictions were randomly assigned to one of the two cells (the minimum restriction) that they opted for; those whose restrictions limited them to three cells were then randomly assigned and finally the unrestricted group were randomly assigned so that the maximum of three classes per cell was attained. After the project got under way, one teacher in the primary, no-aide, increased class received a fellowship and returned to school. Thus, 22 teachers were involved in the project, and it is these classes which provided data in the analysis of the first semester of the study included in a previous progress report.

Because of the problems encountered, certain aspects of the study should be noted. First generalization beyond the present sample of teachers is not really warranted. However, we believe the problems encountered by the investigators were typical of large cities. It does behoove us, however, to describe in considerable detail the teachers and the pupils in the study in order that an interested reader, in the face of a lack of any other information, might be able to generate some educated hunches in relation to his own concerns. Within the design itself, violations of complete randomization suggest that the alpha and beta levels for the two kinds of errors had to be given additional consideration.

Table 2

Restrictions imposed by teachers in becoming part of the project. Check marks indicate the classes to which they were randomly assigned.

Primary	(1) Aid Standard	(2) Aid Increased	(3) No Aid Standard	(4) No Aid Increased
9[£]	✓(1)[££]		✓	
3	✓		✓(3)	
19	✓		✓(3)	
18			✓(3)	
17	✓	✓(2)	✓	
16	✓(1)	✓	✓	
15	✓	✓(2)	✓	
14	✓(1)	✓	✓	✓
4	✓	✓(2)	✓	✓
23	✓	✓	✓	✓(4)
13	✓	✓	✓	✓(4)
Intermediate	**(5)**	**(6)**	**(7)**	**(8)**
10	✓(5)[£££]			
21	✓(5)	✓	✓	
11			✓	✓(8)
6	✓	✓(6)	✓	
7	✓	✓(6)	✓	
5	✓	✓	✓(7)	
2	✓	✓(6)	✓	
22	✓(5)	✓	✓	✓
20	✓(5)	✓	✓	✓
1	✓	✓	✓(7)	✓
8	✓	✓	✓	✓(8)
12	✓	✓	✓	✓(8)

[£] Numerical designation of an alphabetical listing of the teachers.

[££] Bracketed number indicates cell actually assigned to within limits of restriction.

[£££] Just prior to start of first semester this teacher refused to get in project unless assigned an aide. She was accepted.

Prior to the start of the third semester, seven teachers dropped out due to return to school (2), pregnancy (1), husband left town for new job (1), better job offer in town (1), failing health (1), and not adequately certified (1). Four teachers transferred into the project and took over vacated classes. The remaining three classes could not be filled. Thus, 20 classes remained constant within the context of the design and these 20 classes were distributed among 16 schools in the core or near the core area of Milwaukee. Thus, eleven classes were predominantly black, five were predominantly white, and four were mixed (about 50-50).

	Aide		No Aide	
	Standard	Increased	Standard	Increased
Primary	2	3	3	1
Intermediate	4	3	2	2

As noted previously, teachers got into the project with various constraints on their entry. In an attempt to determine if this selection factor has any effect on the dependent measures, a preliminary analysis was done using the WRAT subtests and dependent variables. The interaction between the restricted versus non-restricted interaction with time was the effect of interest. An F-ratio of 1.21 for the multivariate test of equality of the mean vectors was obtained. Since a $p<0.36$ (df = 6 and 13) was associated with this ratio, it was assumed that the manner in which the teachers entered the project has a minimal effect on the dependent measures.

Teacher Orientation

The teachers were involved in a series (not all at the same time) of inservice sessions to (1) further explain the project, particularly the various dependent measures used in the study, (2) explain the use of the Peabody Language Development Kit, which was placed in all classes as a quasi-curriculum control, (3) discuss the functions aides can perform in the EMR classes.

In addition, a three-day conference was held on August 28-30, 1967, in Madison for all participating teachers, aides, project staff, Division for Handicapped Children staff, selected Milwaukee Public School special education administrative staff, and teacher trainees in mental retardation. The program was so structured to review project goals and objectives, give an overview of other viewpoints concerned with aide training and utilization, present concrete examples of aide roles and functions, and give information relative to the possible impact of aides upon the educative structure and process. The teachers and aides were given the opportunity to interact with the project directors and investigators and discuss areas of concern.

Aide Selection

A pool of some 40 aides was readily recruited through word of mouth and more formal advertising. The basic requirements of the aides were that they had to be at least 20 years of age, a resident of Milwaukee, and have had two years of college or two years in a structured setting working with children. Twelve of the aides were randomly assigned to the four cells of the design which required aides, i.e., three to a cell. Replacements were made for the aides who for a variety of reasons left the project. Following are some characteristics of aides who were involved in the project:

(1) They were all female and had had from four to eight years' experience working with children in a structured setting.

(2) They had an average of one and one-half years of college.

(3) Their age ranged from 20 to 46 years with a median of 38 years.

(4) Seven of the 18 aides employed throughout the project were white, the remainder were black.

Aide Training

The original 12 aides received some 40 hours of inservice training (see page 95). The replacement aides received intensive five-days' preservice training encompassing teacher-aide interaction, aide-pupil interaction, role and function, employee considerations, general purpose of project, and specific aspects of project relating to the aide's functioning.

In addition, inservice meetings were held throughout the first half of the school year for all aides. Sessions of two hours' duration were held twice a month. During the second and third semesters, meetings between the project coordinator and aides were held monthly.

Role and Function of Aides

The role and function of the aides grew constantly in both scope of activities and responsibility for activities. Utilization of the aides seemed to vary according to the divisions of teachers and aide competencies, and personalities of the teacher-aide team appeared to dictate the scope and depth of aide activities. During the last year greater use was made of aides as instructional assistants. Aides with adequate academic background under teacher direction developed daily plans for individual and small groups. Greater utilizations were made of the unique strengths of particular aides. For example, those who displayed exceptional sensitivity and management techniques were given greater responsibility in working with youngsters having behavioral adjustment problems.

Listed below are a group of activities performed in many of the classrooms by aides.

(1) Taking attendance - daily
(2) Collect lunch money
(3) Order supplies
(4) Using Duplicator - also the Thermo-Fax
(5) Getting Peabody Kit ready - daily
(6) Running tape recorder
(7) Running film strips
(8) Getting art supplies ready
(9) Fixing bulletin boards
(10) Keeping track of permission slip and money (circus, movies, etc.)
(11) File papers
(12) Check papers
(13) Working with individual students
(14) Giving individual students tasks
(15) Encouraging individual students in personal hygiene
(16) Taking students to the nurse
(17) Supervising class when teacher has to leave the room
(18) Assist with gym class
(19) Cleaning cabinets
(20) Helping plan parties for special events
(21) Help children paste pictures on chart
(22) Setting up easels
(23) Making booklets
(24) Cutting patterns
(25) Write notes for teacher
(26) Took charge of older groups in spelling and reading lessons
(27) Helped on playground
(28) Helped with arithmetic lessons
(29) Collecting and recording collections
(30) Supervising parties and assisting with special occasions - food preparation, games, etc.
(31) Make monthly calendar
(32) Clean paint and paste brushes
(33) Check workbooks
(34) Make seating chart for substitute
(35) Water plants
(36) Repair toys
(37) Get and return books from cupboard
(38) Assisting children with outer wraps
(39) Walk children to bathroom and for recess
(40) Fill out form for movies for Milwaukee Public Library
(41) Line chart paper for reading readiness
(42) Make color chart
(43) Helping plan and helping supervise cooking lessons - that they can develop concepts of measuring liquids and solids - also planning meals and use of money
(44) Distribute working materials - paper, workbooks, etc.
(45) Helping children to get settled for their daily tasks
(46) Getting children to dress for recess and for going home

Selection of Pupils

Since the teacher was the experimental unit in the study, the children in the respective classes constituted the major portion of the study sample. Children assigned to the increased classes were taken from the waiting lists of the Milwaukee Public Schools under one restriction. In anticipation of pupil losses over the project, an attempt was made to have as much consistency of pupil population as possible by requiring a maximum age of eight for assignment to an increased primary class and a maximum of eleven for the increased intermediate classes. In the Milwaukee system, pupils move from primary to intermediate classes around ten years of age and from intermediate to secondary classes about age thirteen.

The number of pupils involved in the project at the outset (when there were 23 classes) was:

	Aides		No Aides		
	Standard	Increased	Standard	Increased	Total
Primary	(3) 35	(3) 50	(3) 32	(2) 34	151
Intermediate	(4) 53	(3) 63	(2) 29	(3) 58	203
Total	88	113	61	92	354

The overall number of pupils involved in the project remained quite constant (about 354) during the first semester (February - June, 1967). The loss of three classes prior to the start of the fall semester (September, 1967) reduced the overall number of pupils involved accordingly. However, this reduced number (about 306) remained fairly constant over the 1967-68 school year. This is not to say that the exact same children remained throughout but rather that the overall count was maintained by replacement. Indeed of the original 354 pupils in the project, only 232 remained in the project after the three semesters.

Differences among the cells of the design in terms of CA, MA, and IQ were as follows:

		Aides						No Aides					
		Standard			Increased			Standard			Increased		
		CA	MA	IQ	CA	MA	IQ	CA	MA	IQ	CA	MA	IQ
Primary	X̄	92.1	63.1	68.6	94.6	63.8	67.6	92.2	64.3	68.9	101.1	67.4	66.8
	SD	10.7	8.6	5.3	10.8	10.0	7.2	5.8	6.8	6.4	2.9	6.4	7.1
	N		11			19			17			13	
Intermediate	X̄	129.5	89.3	68.0	124.5	84.1	66.5	126.2	86.2	68.3	138.5	94.2	67.8
	SD	10.8	13.7	8.1	10.3	10.5	6.7	10.8	10.8	6.1	7.3	12.3	6.8
	N		23			34			13			19	

Analysis of variance of CA obtained an F-ratio of 70.30 significant at a $p<.001$ ($df = 7, 143$). This difference was expected as CA was a distinguishing characteristic of the primary and intermediate groups. No significant difference was obtained among the cells in terms of IQ ($F = 0.38$, $p = .911$, $df = 7, 143$). By definition, since CA was different and IQ not, we would expect significant differences among the cells in terms of MA. This was the case ($F = 25.7$, $p<.001$, $df = 7, 143$). The logic here is that if IQ is a constant (which it was) in the formula $IQ = \frac{MA}{CA}$ then to maintain the constant, MA would have to vary in the same direction with roughly equal magnitude as the CA varied among the cells.

One final point, the N's noted in the above table were the number of pupils who were both in the project and given the WRAT on all three occasions. Contrast this with the previous table where the count is in terms of absolute numbers in the project at the beginning.

B. Dependent Measures

Pupil change in three basic areas constituted the bulk of the dependent measures in this study. The three areas referred to are academic, language, and behavior. The instrument used for measuring change in academic skills was the Wide Range Achievement Test (WRAT) (Jastak and Jastak, 1965). In language it was the Illinois Test of Psycholinguistic Abilities (ITPA), developed by James J. McCarthy and Samuel Kirk (1961), and tape-recorded samples of individual pupils' speech. In behavior it was the Behavior Analysis Scale - Educable (BASE), indices of behavior from the longitudinal films and a measure of absentee rate.

WRAT

The WRAT measures competency in reading, spelling, and arithmetic. It has items in it which have been standardized on pre-kindergarten children; thus, (with only one or two exceptions) all of the children in this study were able to pass some items each time they took the test. The test was administered individually to all children in the project in the early part of the spring semester of the 1966-67 school year, and in the early part of the fall semester of the 1967-68 school year. In the latter part of the spring semester of the 1967-68 school year this test was readministered to all students still in the project to whom it had previously been administered at either, or both, the preceding spring and fall testing.

The WRAT was administered by personnel especially trained for this purpose. After the first semester, these testers consisted largely of the aides from the project. Otherwise they were personnel employed specifically for this purpose, or who were ordinarily performing other functions on the project but who were used as testers during the time that testing was the high priority activity.

ITPA

The Illinois Test of Psycholinguistic Abilities (ITPA) indexes language development by providing measures of linguistic attainment in the following areas:

(1) Decoding, which refers to those abilities required to obtain meaning from visual and auditory linguistic stimuli, i.e., receptive language ability.

(2) Association, which is the ability required to manipulate linguistic symbols internally. This is a central process elicited by decoding which in turn elicits expressive processes.

(3) Encoding, which is the sum of those abilities required to express ideas in words and gestures.

The test measures a further breakdown of these processes in the sense that it provides a measure of the child's functioning with respect to auditory input, vocal output, visual input, and motor output.

In the first semester of the study (spring, 1966-67) the ITPA was administered to a sample of pupils (ten students from each class) who appeared to be most likely to remain in the class for the duration of the study. The second, or post, administration of the test was during the latter part of the spring semester, 1967-68 school year (the third semester of the study). At that time it was re-administered to all students still in the project who had had the test during the first semester.

During the first testing with the ITPA, children were bused into a centrally located school on four consecutive Saturdays. Testing was done by school psychologists from the Milwaukee School System. During the week, especially hired and trained testers visited classes and picked up those children who were unable to be at the Saturday sessions. The second testing (in May, 1968) was done by a specially marshalled and trained force of testers. These were augmented by a bus, in which were several testing booths, which could be used at those schools where space for testing was a problem. These people devoted approximately two weeks to this task.

Speech Samples

Speech samples were three minutes (or approximately so) of individually tape-recorded speech of the children. These were obtained by presenting two Charles Addams drawings sequentially to the subjects and asking the subjects to talk about these drawings (see Appendix C for a sample of instructions). Different sets of drawings were used each of the three times (or testing periods) the speech samples were taken. Later, through analysis of these samples by trained linguists, a number of indices speculated to be sensitive measures of change in language proficiency were obtained. These were: number of words per minute, mean words per narrative segment, average number of syllables per token, and various ratios between a number of raw indices (see Appendix C for a complete listing of both raw and derived indices).

In the first two testing periods (early part of spring semester, 1966-6 and early part of fall semester, 1967-68) speech samples were obtained on all students in the project. In the third testing period (latter part of spring semester, 1967-68) they were obtained only on 80 children. It was decided because of economics and time to concentrate only on a random sample of ten students from each of the eight cells of the study design. Hence, in the third testing period, speech samples were obtained only on those 80 children earlier selected to make up this sample.

As previously stated (in the paragraphs with reference to the ITPA), during the first testing period some speech samples were taken by testing psychologists at the time they were giving ITPA's to the children, and some were obtained during these same special Saturday sessions by testers specially trained for this activity. Children not sampled on these early Saturdays were later sampled by these same, and other, specially trained testers, either in conjunction with testing with the WRAT or on separate occasions, whichever fitted the individual circumstance.

BASE

The Behavior Analysis Scale - Educable (BASE) is a 70-item scale on which the child is rated by an adult who is in a position to know and to observe that child. This instrument (see Appendix D for test and directions was developed in cooperation with the University of Wisconsin's Department of Counseling and Behavioral Studies. Before use in the study, the instrument had been field tested on some 400 EMR students in the state. Item analysis reduced the original 130 items to the 70 which made up the instrum as it was used in this study.

Previous work with the scale indicated it would provide five stable factors or clusters which could be meaningfully labeled and used as dependent variables in the present study. That is, it was anticipated that scores derived from these factors at different times during the study would reflect behavior changes occurring in the classroom over the period of the project.

All children in the study at the time were rated on the BASE during the first and second testing periods (spring, 1966-67, and fall, 1967-68). Only those children who had previously been rated during th first and/or second testing periods were rated in the third testing period. A sample rating was also done a few weeks after the first rating to determine the reliability of the instrument.

Raters of the children on this instrument were the teachers in the respective classrooms. In those classes where aides were also present, the aides also rated the children at the same time as the teachers. Only the data from the teachers' rating were used in the statistical tests for effect of treatments. Ratings made by the aides have not been used at this time.

The two ratings done in the early parts of their respective semesters (testing periods one and two) were deliberately done several weeks after the start of the semester. This was in order to give teachers (and aides) that degree of experience with the pupils (particularly any of the pupils who may just have entered the class at the beginning of that semester) before requiring them to rate their children. It required about ten minutes to rate a pupil.

Expectancy Index

A tangential concern in the study was the relationship, if any, between the expectancy on the part of classroom adults, or improvement in a child, and that child's actual improvement, over a period of time, in any of the considered areas of ability or skill. In order to incorporate this concern into the study, a simple expectancy index was developed on which raters could index their expectancy of a given child's improvement in the realms of social, language, and academic behavior. Appendix A contains an article based on this data.

Audio-Visual Monitoring

As originally conceived, audio-visual monitoring of the study was to be done using video tape. Problems of personnel, in terms of numbers required and qualifications, and on-site difficulties such as electrical supply, physical plant and classroom arrangement forced the use of films instead. Permission was obtained from the Grants Officer to make this change. We were fortunate to be able to hire a competent audio-visual person with considerable experience in classroom filming. By April, 1967, we eventually obtained the filming equipment that had been ordered eight months previously. It was possible, however, to get the shooting schedule under way by using loaned equipment from the supply company.

Two skilled professionals advised in the area of audio-visual monitoring. Mr. Walter Meives, Director of the University Photo Lab, helped in the technical and production end of the filming. Dr. Robert Boyd of the University of Wisconsin's Department of Curriculum and Instruction consulted regarding the development of film-derived indices appropriate to the goals of the study and the population with which we were working. These two men represented, in a sense, somewhat opposing positions, both of which were a legitimate concern and both of which we were trying to incorporate into our production. That is, on the one hand we had a concern with a sensitive, artistically-inspired and flawless production, and, on the other hand, we had a concern with the production of observable data which was meaningful in a research sense.

In spite of the fact that we were shooting without special lighting, the black and white prints came out remarkably clear. A few problems occurred with the sound in earlier shots because of the natural setting, but they were solved fairly effectively.

The filming schedule was divided into four time periods:

Period 1 - morning, before recess
" 2 - morning, after recess
" 3 - afternoon, before recess
" 4 - afternoon, after recess

Within a given time period, five minutes of film were shot about ten minutes after the beginning of the period and about ten minutes before the end of the period. In this way curriculum changes within a period were captured. A block of filming was considered the filming of all the classes in the project.

The original plan for the filming was that each class would be shot four times (e.g., there would be four blocks of filming) during the course of each semester, recycling the classes among the periods so a complete day's activities could be caught. However, after the first two blocks it became apparent that as a result of (1) cost of filming, both in terms of money and classroom disruption, (2) limitations on the rate at which the needed film was becoming available, and (3) a closer look at the actual needs of the project with regard to the filming, continuing with the plan of four blocks per semester no longer seemed reasonable. Therefore, Block III for the first semester was cancelled. Also, during the 1967-68 school year, the number of blocks over the two semesters was reduced to four, evenly spaced throughout the year.

Thus, during the first semester of the project, 30 minutes of film on each class, or 11 1/2 hours in all, was shot; during each of the second and third semesters, 20 minutes was shot for each class. A total of 70 minutes of film was obtained on each of the 23 classes. Additional footage was spent working bugs out of the equipment and capturing those activities such as aspects of the testing procedure, use of the Peabody Language Kit Materials, utilization of teacher aides in various settings, field trips, and interpersonal interviews which we felt would enhance the final production.

Noise Level

Noise level in the classroom, and its relationship to or effect on, learning was one of the secondary concerns in this study. Measures of noise level were taken concurrently with the film shots. Equipment used for these measurements were a Nagra III Recorder and Sennheiser Microphone. These were placed at various locations in the rooms. The time involved was no shorter than ten seconds nor longer than three minutes. Measures for each filmed segment included the peak noise level or loudest noise observed, the average volume of the teacher's and aide's voices, and the ambient level defined as the noise level of the room excluding individually distinguished sources other than continuous mechanical noises. Appendix A contains an analysis and write-up of this data.

Film Indices of Behavior

After initially viewing the films, the question arose of just how much the cameraman's presence was interfering with the class activity. Were we getting a representative example of what was going on in these particular classes day after day, or was the presence of the camera disrupting normal processes? At one extreme, the pupils seemed to be putting on a show for the benefit of the cameraman. At the other, they seemed to be almost completely oblivious to his presence. In order to get an indication of the amount of interference taking place, we counted the number of students gazing at, or performing for the camera within arbitrarily selected units of time - in this case, every three seconds. Data sheets with each ten-minute filming sequence divided into three-second intervals were drawn up, and the number of children actually looking at the camera within these units of time was recorded. There was a problem with this method of recording distractions related to the filming technique. The cameraman focused on individual children for the most part, and scanned the entire class relatively infrequently. When the camera moved away from a child who was staring intently at it, the child ceased to be recorded. Yet an observer could guess, it seemed quite accurately, that the child was still staring, especially if the camera re-focused on him a short time later to find him similarly engaged. Only the observable behavior was recorded, so that the total count of camera distractions was less than it would have been had the entire class been in focus. The cameraman's movement around the classroom attracted additional attention. However, the use of a stationary camera would have resulted in the loss of other types of information: the activity of a group in the far corner of the classroom, the activity of individuals involved in independent work, and the type (and often content) of exchanges between members of the class.

When it was found that the filming sequences usually ran over ten minutes, each sequence was timed and the number of three-second intervals adjusted accordingly. The length of the two segments into which each film was divided varied, but each of these segments was timed and the number of intervals adjusted.

It was found that the number of children distracted by the camera decreased radically from one five-minute segment to the next. The same results were not obtained for successive blocks, however. The number of distractions seemed to depend to a large extent on the type of activity taking place. There were more classes whose distraction count decreased rather than increased, however, (16 as compared to 11) over successive blocks. It was not a steady decline from one block to the next - the count would often jump or decrease as much as 40 points between blocks.

A count was obtained from the films of the number of separate instances of children engaged in non-task oriented behavior (NTO). Excluded were all distractions connected with the camera. The number of children who, at distinct instances, were doing something other than the task at hand were counted. We divided this behavior arbitrarily into categories of restlessness, wandering, socializing, daydreaming, and other activity. In restlessness were included all those pupils who were seen bobbing up and down in their seats, playing with pencils and materials. These activities had to be accompanied with not paying attention. If the child was restless, yet involved in the lesson, he was not counted. In wandering were included all the pupils who seemed to be wandering aimlessly. It was difficult to determine in some cases, since a child would walk across the field in view and his destination would be unknown. If his wandering seemed to be connected with the task (this was especially evident when a class was engaged in arts and crafts), he was not counted. In socializing the same held true. If a child seemed to be talking about the activity he was engaged in, he wasn't counted. In daydreaming were included those kids who were staring out a window or sleeping. If a child was staring at an activity going on in another part of the room, however, he was included under another activity, rather than daydreaming. These activities had to be separate and distinct. If the camera moved away from a child that was socializing for only a very short time, the child was not counted twice.

The count of non-task oriented behavior (restlessness, wandering, socializing, daydreaming, and other activities combined) was taken for each film in three-second intervals. This information was integrated with the camera distraction count, which was also taken in three-second intervals. This gave a general index of the amount of inattentiveness or restlessness in each class or as it was termed, activity level (AL).

In sum, the films yielded for analysis three measures of classroom behavior: non-task oriented behavior (NTO), activity level (AL), and noise level (NL).

Light Levels in Classrooms

Light readings were routinely collected during the course of the filming sequences. A descriptive analysis of this data is provided in Appendix A.

Absentee Rate

Absentee rate for each pupil for each of the three semesters of the project was calculated from information obtained directly from the teachers or by staff personnel from attendance records maintained in the offices of the respective schools. The percentage of days absent was a function of the total possible number of days present. Ordinarily the school year is 180 days, but sometimes mitigating factors such as return to regular class, sent to boys' school, or moved out of area reduced the number of days.

Meteorological Indices

It has been the observation of those involved in day-to-day classroom programming that meteorological factors may have a considerable effect upon the pupil's behaviors. At the start of this project no attempts, to our knowledge, had been made to document this observation.

The filming schedule for this project resulted in the sampling of classroom activity over a considerable range of time and across a wide variety of classes. With a minimum amount of additional effort (on the part of our cameraman) we also collected measures of variations in barometric pressure, outside and inside temperature, and wind speed. These were collected for correlation with indices of noise level, activity level, etc., which the concomitant film segments provided. Preliminary analysis has been done but not written up for this report.

Scoring and Additional Data

Scoring on all tests, scales, and measures was carefully checked and reviewed for accuracy and consistency, and analysis of speech samples and filmed segments was carried out by personnel at the headquarters for the project in Madison. Appendix D contains detailed discussion of the modifications in the testing and scoring procedures.

Addresses, birthdates, and IQ's on all pupils in the project were basically obtained from copies of forms used by the Milwaukee Public Schools for providing the Division for Handicapped Children, Wisconsin Department of Public Instruction, with EMR class enrollment information. Information on the race and sex of each of the children in the project was obtained by specifically requesting it from the teachers, either as an isolated item or in conjunction with some one of the other forms, particularly the BASE, which they filled out and provided on each of the children.

Survey of Interpersonal Values

The only test given to teachers and aides themselves was the Science Research Associates, Incorporated (SRA) Survey of Interpersonal Values (1960). The purpose of this test was to help characterize those people who work well together in the particular situation in which teachers and aides were working in this project.

As the title indicates, this test measures one's interpersonal values. in this context, values should be differentiated from personality traits which in essence describe what a person does characteristically in particular situations. Values, on the other hand, describe the person in terms of his basic emotional patterns.

The six values measured by this test are:

(1) Support
(2) Conformity
(3) Recommendation
(4) Independence
(5) Benevolence
(6) Leadership

With the exception of a sampling of teachers and aides who took the test twice in order to provide a reliability check, all others took it only once. This was sometime during the first two or three months after the time they actually began to function in the classroom as part of this project. Those people who took the test twice were people who were in the project from the start. They took it for the first time within two or three months of the beginning of the first semester and for the second time within two or three months of the beginning of the second semester.

Directions for the test are self-explanatory and contained on the front page of the test form. Thus, although the test can be administered in a group, for our purposes each teacher and aide was given the test to do on his own.

Analysis of this test consisted of rating the teacher-aide dyads in terms of how well they worked together on a three-point scale. A profile analysis revealed nothing of significance and further use of the test was not considered.

CHAPTER V

RESULTS AND DISCUSSION

A. Quantitative[1]

This chapter contains two broad subdivisions of the results - quantitative and qualitative. The former division consists of analyses of the hard numerical data of the study while the latter division presents the subjective reactions of participants in the project. The quantitative division in turn consists of three sections. The first deals with the analyses of the various dependent measures of the study; the second section attempts to define the characteristics of those pupils which demonstrated the greatest academic gain over the three semesters of the project; the third section presents a cost achievement analysis to assist administrative decision making.

1. Analyses of Dependent Measures

a. Academic: WRAT Raw Scores

Observed mean scores for the WRAT subtests of reading, spelling, and arithmetic were obtained for each of the 20 classes in the study at three points in time, i.e., February, 1967, September, 1967, and May, 1968. The first objective was to determine if these scores - when averaged over the three time periods - displayed significant variation by type of class (primary or intermediate) presence or absence of aides, size of classes (standard size or 50% increase in size) or any interactions among these factors. The second objective was to determine if these scores changed with time and if the class type, aide, size, or interactions among these had an effect on the scores over time.

Preliminary Analysis

To better study any trend of change with time, orthogonal polynomials transformations were used to isolate linear and non-linear components of the change. Since the interpretation of effects and interactions is greatly simplified if the non-linear components of the time trend can be regarded as insignificant (both statistically and otherwise), the first test was a test of the composite hypothesis that non-linear component of the time trend and all interactions with the other factors was zero for all three dependent scores. This test used the sum of squares combined from all sources of variation with eight degrees of freedom. The results of this analysis are shown in Table 3.

---

[1] This section of this chapter was written in conjunction with Tom Fischbach, statistical consultant with the Research and Development Center in Madison Wisconsin.

Table 3

Non-Linear Component of Time Trend for WRAT Scores

Variable	Hyp. of No Effects SS	df	Hyp. MS	Error MS	Error df	F ratio	Sig. Level
Reading	14.96	8	1.87	3.23	12	0.58	p<.900
Spelling	22.77	8	2.85	2.85	12	1.00	p<.500
Arithmetic	12.53	8	1.57	0.37	12	4.24	p<.025

To reject the null hypothesis for this analysis the F-ratio for any dependent variable should be significant at the 0.15 level rather than the customary 0.05 level since failure to reject the null hypothesis when a real non-linear trend exists would result in an over-simplification of the true state of affairs. With this noted, it is clear that the non-linear trend component effects were statistically insignificant in the case of reading and spelling. But this was not the case for arithmetic. This was true even when nominal significance at the .05 level was required (one-third of .15) to allow for the fact that the three dependent variables were correlated.

Further inspection of the data revealed that of the total sum of squares of 12.53 for non-linear trend for arithmetic, most of it (10.62) was for the main effect so that only 1.91 of the total was for the interaction of the non-linear time component with other factors. Using Scheffe's S-method of simultaneous inference, (Scheffe, 1959) that is, treating the main effect component as though it had eight degrees of freedom instead of one gave a mean square of 10.62/8 or 1.33 for arithmetic. This effect was significant at the 0.025 level as the F-ratio was 3.59.

Since the linear component of the time trend for arithmetic was actually confounded with non-linear components, it was necessary to combine the two components in the analysis of the effects of time interactions with other factors on arithmetic scores. This was not necessary for reading and spelling.

Analysis of Variance: Between Cells

The results of the analysis of variance are shown in Table 4 (a,b and c). This table is divided into three parts, one for each dependent variable. The first half of each of these parts shows the results of the analysis of between class variations by condition while the second half shows the results for the "within class" over time analysis.

Interactions

The first hypothesis to be considered was the null hypothesis of no effects of interactions among the three conditions or characteristics of the classes on the dependent variables. In all cases the data does not require rejection of this hypothesis. This was also substantiated by a multivariate analysis entering all three WRAT scores in the analysis. Thus, only main effects needed to be considered.

Class Type

For all three dependent variables it was necessary to reject the null hypothesis of no difference between primary and intermediate classes. The estimates of the contrasts of the mean for primary classes with intermediate classes averaged over all three time periods is shown in Table 3 (a, b and c). In all cases the means for intermediate classes were higher than those for primary classes.

Aides

In no case was it necessary to reject the null hypothesis that the presence or absence of aides did not affect class mean scores. However, for all three variables the mean scores for classes with aides was higher than for the classes without aides as may be seen in Table 3. But, the effect of the presence of aides must be considered statistically insignificant in light of this data.

Size

The null hypothesis that the size of the class did not affect scores cannot be rejected with this data. The mean scores for standard classes were higher than those for increased classes for reading and spelling, but the reverse was true for arithmetic. However, all differences must be regarded as small and statistically insignificant.

Discussion

All of the conclusions previously stated were substantiated by multivariate analyses considering all three dependent variables together. Moreover, reversals of the orders in which effects were removed from the variable did not change the main results. Thus, the major source of variation was due to class type with the intermediate classes having higher scores. This finding was hardly startling since on an a priori basis one would have expected the older (intermediate) children to achieve academically at a higher level on the average over the three testing times.

Analysis of Variance: Within Classes

Interactions of Time with Other Factors

To simplify the analysis the first hypothesis considered was that the effects of all interactions of time with class characteristics were zero. In no case was it necessary to reject the hypothesis. Thus, only main effects needed to be considered.

Time

For all three dependent variables the hypotheses of no change in scores over time were rejected. For reading and spelling, only the linear time trend was studied as the preliminary analysis permitted exclusion of non-linear components. In the case of arithmetic scores both components of time were included and a separate analysis using Scheffe's S-method was used to evaluate the significance of the linear component, which was significant. In all cases the estimates of the linear trend shown in Table 3 indicate that scores increased in each of the two periods. The fact that only the linear component, for reading and spelling, was significant indicates that the increase was uniform over the period of the study. The "quadratic" component for arithmetic was also positive indicating a greater change during the second period.

Class Type x Time

The null hypothesis that the change over time for primary classes was the same as for intermediate classes had to be rejected for all three dependent variables (see Table 4 a,b, and c). The differences between the two types of classes - in favor of intermediate classes - decreased with time as the estimate of the class x linear time trend effect for all variables was positive, indicating that the slope of the linear trend was greater for primary classes than for intermediate classes (see Table 5, b). The preliminary analysis indicated that the interactions of the non-linear time trend components with other factors was not significant for the case of arithmetic. However, in all cases, the estimate of the quadratic component was also positive, indicating a slight tendency for the relative increase over time of the slope of the primary classes.

Aides x Time

The null hypothesis that the increase in scores with time was no greater for classes with aides could not be rejected with this data for any variable. However, the estimates of this effect are positive for all three variables which indicates a slight tendency for a greater increase over time for classes with aides (see Table 5, b). However, these differences could not be regarded as statistically reliable indications of true differences; any real differences were small.

Size x Time

The null hypothesis that the change over time for standard classes differed from increased classes could not be rejected for reading or spelling scores but had to be rejected for arithmetic scores (in this case) because only one of the three cases is nominally statistically significant, the criterion was that the F-ratio must be significant at the .05/3 or .0167 level; which was obtained for arithmetic scores (see Table 4,c). For all three variables the estimate of the size x time interaction was positive indicating a tendency for the slope of the line to be greater for the standard sized classes. But, only in the cases of arithmetic scores could these differences be regarded as statistically reliable estimates of true differences (see Table 5, c).

Table 4

Analysis of Variance for WRAT Scores

a) Reading

Source (1)	Sum of Squares	df	Mean Square	F-ratio	Significance Level
Between Classes	—	20	—	—	—
Mean	—	1	—	—	—
Class Types	2093.32	1	2093.32	54.12	p<.001
Aides	5.53	1	5.53	0.14	p<.710
Size	91.66	1	91.66	2.37	p<.150
All Interactions	132.13	4	33.03	0.85	p<.730
Error (Within Cell)	464.21	12	38.68	—	—
Within Classes by Time (Linear Only)	—	20	—	—	—
Time (Linear)	512.44	1	512.44	106.24	p<.001
Class Type x Time	61.93	1	61.93	12.83	p<.004
Aides x Time	3.27	1	3.27	0.68	p<.430
Size x Time	2.25	1	2.25	0.47	p<.510
All Other Interactions	35.10	4	8.78	1.82	p<.300
Error (Class x Time Within Cells)	57.91	12	4.83	—	—

b) Spelling

Source	Sum of Squares	df	Mean Square	F-ratio	Significance Level
Between Classes	—	20	—	—	—
Mean	—	1	—	—	—
Class Types	578.32	1	578.36	60.47	p<.001
Aides	0.86	1	0.86	0.09	p<.770
Size	1.07	1	1.07	0.11	p<.760
All Interactions	23.48	4	5.87	0.61	p<.750
Error (Within Cells)	114.81	12	9.57	—	—
Within Classes by Time (Linear Only)	—	—	—	—	—
Time (Linear)	139.50	1	139.50	92.94	p<.001
Class Type x Time	14.27	1	14.27	9.51	p<.009
Aides x Time	2.16	1	2.16	1.44	p<.250
Size x Time	0.44	1	0.44	0.29	p<.600
All Other Interactions	1.02	4	0.25	0.09	p<.900
Error (Class x Time Within Cells)	34.17	12	2.85	—	—

Table 4 (Cont.)

Analysis of Variance for WRAT Scores

c) Arithmetic

Source	Sum of Squares	df	Mean Square	F-ratio	Significance Level
Between Classes - Total	—	20	—	—	—
Mean	—	1	—	—	—
Class Types	562.56	1	562.56	73.85	p<.001
Aides	0.91	1	0.91	0.12	p<.740
Size	4.13	1	4.13	0.54	p<.480
All Interactions	69.73	4	17.43	2.29	p<.100
Error (Between Classes/ Cells)	91.40	12	7.62	—	—
Within Classes by Time	—	40	—	—	—
Time	218.10	2	109.05	336.11	p<.001
Class Type x Time	10.29	2	5.15	15.87	p<.001
Aides x Time	0.02	2	0.01	0.04	p<.930
Size x Time	4.32	2	2.16	6.65	p<.010
All Interactions	5.77	8	0.72	2.22	p<.100
Error (Class x Time Within Cells)	7.79	24	0.32	—	—
"Linear" Components (ii)					
Time	204.48	1	204.48	743.22	p<.001
Class x Time	10.11	1	10.11	36.72	p<.001
Size x Time	4.27	1	4.27	15.28	p<.050
Class x Time, Within Cells	3.30	12	0.28	—	—

(i) Effects were removed in the order shown. Reversals in order did not substantially change results.

(ii) Scheffe's S-method was used to evaluate significance, e.g., each F-ratio shown in the table was divided by 2 and referred to a table with 2 and 24 degrees of freedom.

Table 5

Estimates of Effects

Effect	Estimate			Standard Errors		
	Estimate and Standard Error of Estimate:					
	Reading	Spelling	Arith-metic	Reading	Spelling	Arith-metic
a) Intermediate-Primary	20.08*	10.44*	10.03*	3.01	1.50	1.36
Aides vs. No Aides	1.22	0.31	0.19	3.01	1.50	1.36
Standard vs. Increased	3.52	0.11	-1.14	3.01	1.50	1.36
Time:						
Linear	4.77*	2.57*	3.23*	0.53	0.50	0.13
Quadratic	0.81	1.08	0.80*	0.44	0.41	0.13
b) Differences in Linear Slopes						
Primary vs. Intermediate	3.16*	1.62*	1.49*	1.06	0.59	0.26
Aides vs. No Aides	1.26	0.80	0.29	1.06	0.59	0.26
Standard vs. Increased	0.98	0.37	0.92*	1.06	0.59	0.26
c) Differences in Change of Linear Slope:						
Primary vs. Intermediate	0.20	0.70	0.26	0.87	0.82	0.29
Aides vs. No Aides	-0.01	0.05	-0.03	0.87	0.82	0.29
Standard vs. Increased	-0.42	-0.38	-0.16	0.87	0.82	0.29

*Null hypothesis that the true value was zero has been rejected in the analysis.

Summary

The results described for time trends and variations in time trends by class characteristics were substantiated by multivariate analyses. In addition, reversals of orders of removal of effects did not markedly alter the results. The main conclusion was that scores increased with time during the two periods ($T_1 - T_2$ and $T_2 - T_3$) studied and the arithmetic scores increased significantly more in the second period than the first (see Figures 1, 2, and 3). The increase was greater for primary classes than for intermediate classes (see Figures 1, 2, and 3) while arithmetic scores for standard classes also increased more with time than did the scores for increased size classes (see Figure 3). The side effect over time for the three dependent variables is shown in Figure 4. The essential parallelism of the side-on side effect over time reflects the results of the statistical analysis.

Because of the unequal N's in the cells, the least square estimates of the means based on the statistical analysis should vary somewhat from the raw means shown in the graphs. Comparison of these data had indicated the differences were minimal as far as general trend and even size was concerned.

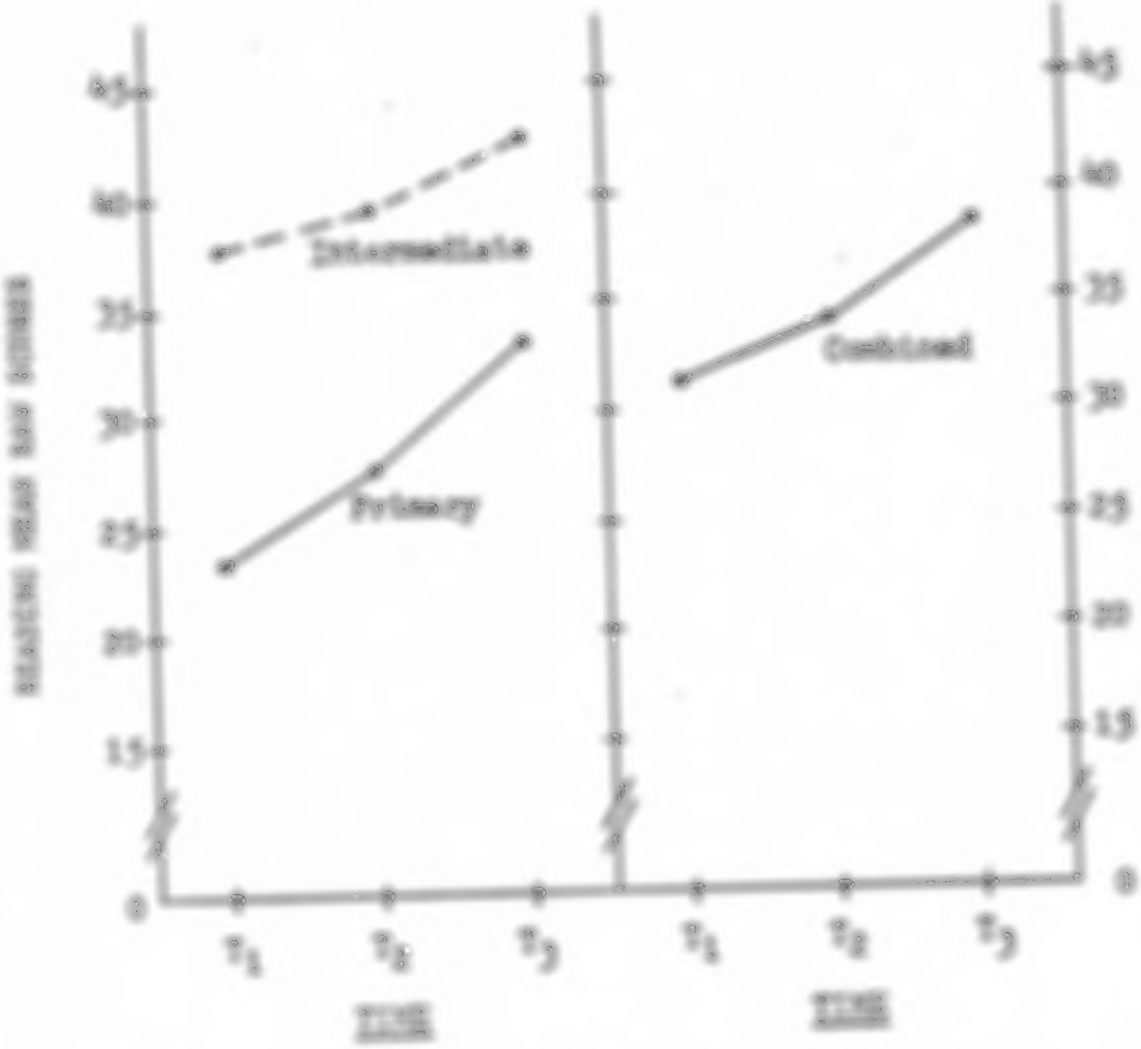


Fig. 1 - Reading scores on the WRAT as a function of class for IDS pupils over three semesters.

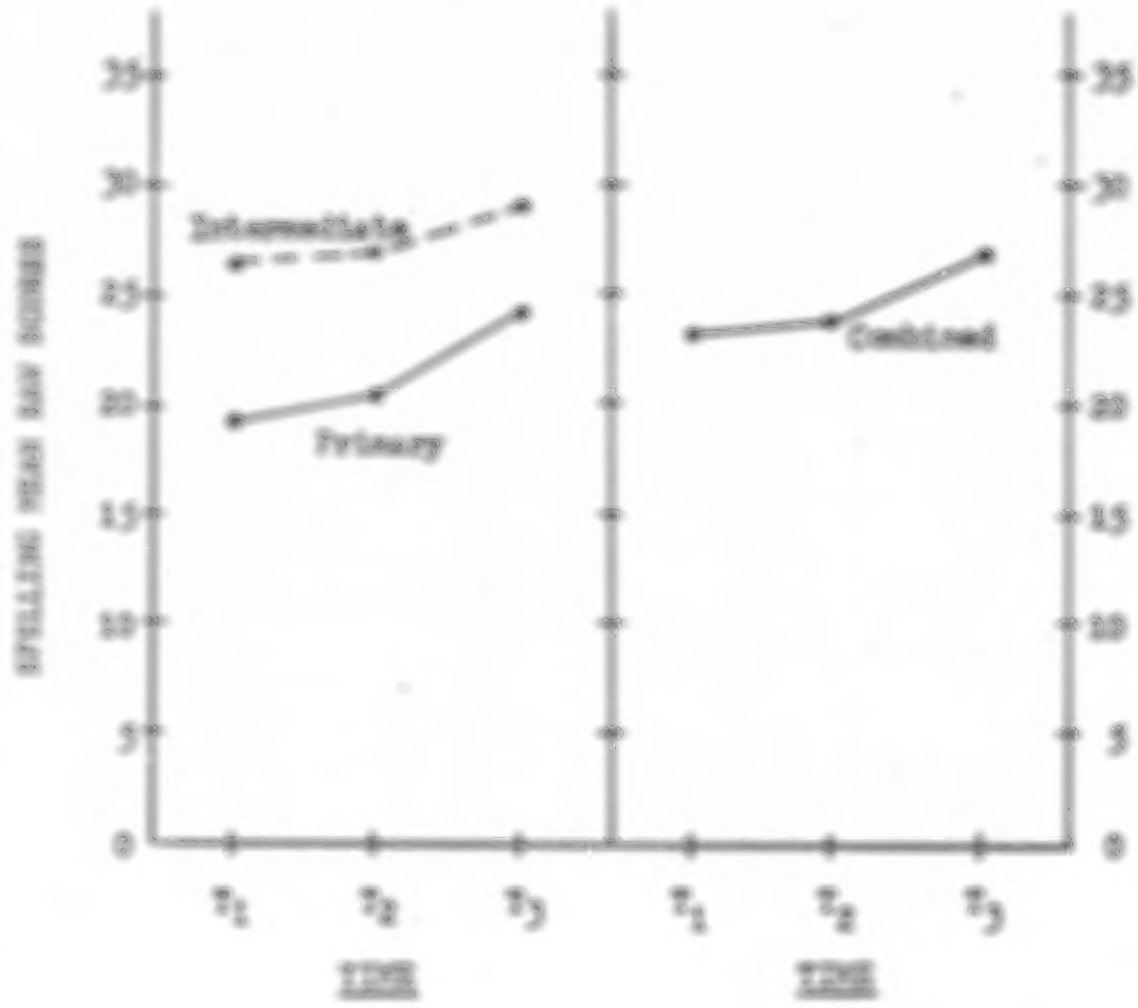


Fig. 3- Spelling scores on the WRAT as a function of class for EMR pupils over three semesters.

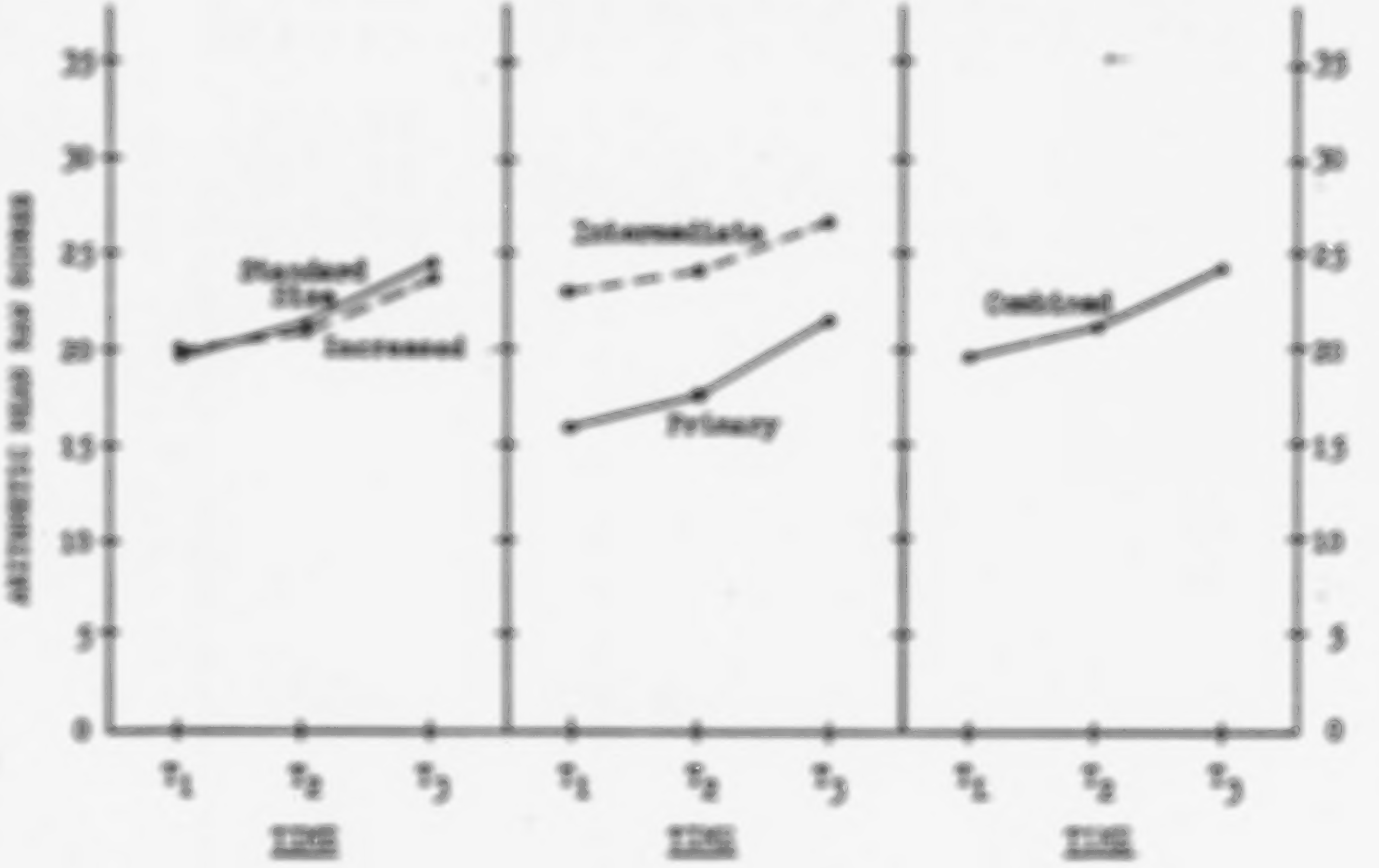


Fig. 3- Interaction of arithmetic scores on the WRAT as a function of class and class size for EMR pupils over three semesters.

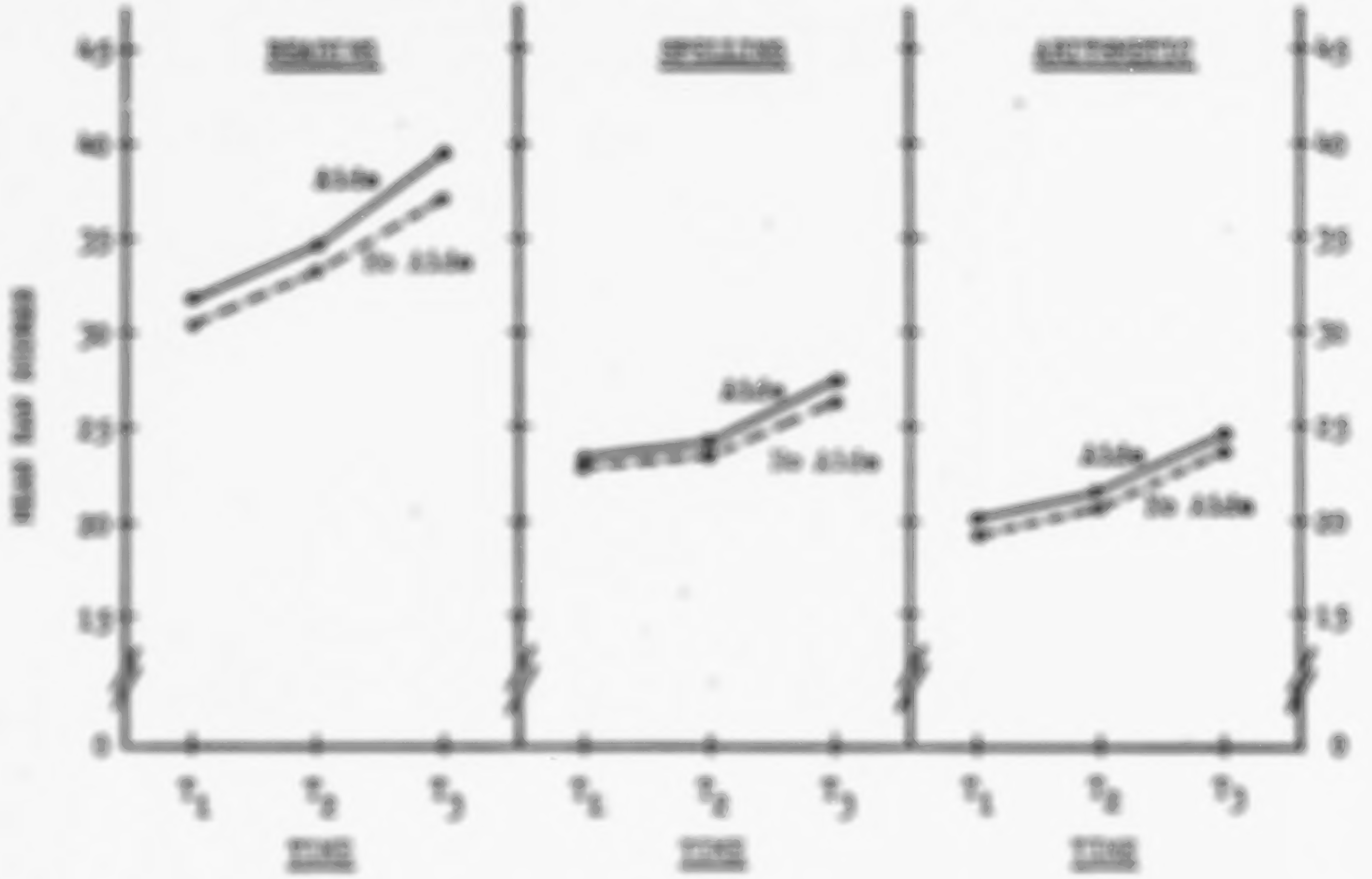


Fig. 4 — Reading, Spelling, and Arithmetic scores on the WRAT as a function of Aides for 533 pupils over three semesters.

b. Behavior Indices

A number of indices of behavior were hypothesized to be differentially affected by the treatment effects. These dependent variables were the five subtests of the Behavior Analysis Scale - Educable (BASE) developed for the project and labeled Disruptive Behavior (1), Attentive Involvement (2), Timidity (3), Dependence (4), and Withdrawal (5). The numbers in parentheses provide an easy reference for the five scales. Additional behavior indices derived from the filming were Noise Level (NL), Activity Level (AL), and Non-Task Oriented Behavior (NTOB) as previously defined. The final behavioral index was that of Absence Rate, the ratio of unexplained absences to legitimate days expected in school over the three semesters of the study.

Data Analysis

Four of the five scales (1, 3, 4, 5) had been developed in such a way that a higher score indicated less desirable behavior (see Appendix B). With scale 2, Attentive Involvement, however, a higher score indicated more desirable behavior. Accordingly scale 2 was multiplied by -1 so that its scores would have the same direction as the others. As may be seen from Table 6, the five scales did have positive correlations with each other.

Table 6

Correlations Among Average BASE Scores for the Three Test Points (Elements Above Diagonal Omitted)

BASE Scale	1	2	3	4	5
1	1.00				
2	0.23	1.00			
3	0.50	0.32	1.00		
4	0.70	0.27	0.72	1.00	
5	0.64	0.33	0.32	0.69	1.00

$df = 12$

For even 12 degrees of freedom the correlations must be considered at least moderate since all were positive. This suggested that a multivariate analysis of variance would be the proper manner to investigate the sources of variation which were of principal interest; the extent to which changes over time on the indices varied with class type, the presence or absence of aides, or with standard or increased size classes, or their interactions. As may be verified by inspection of Table 7, the five subtests varied by scale factors. Scales 3 and 4 had the lowest variance while scale 1 had the highest variance. Thus, it was unwise to restrict the analysis to an analysis of the mean of the five scales only. Hence, a five-variate analysis of variance was required to adequately explain the variance.

Table 7

Variance of SAIS Scale Scores (Average of Three Test Points)

SAIS	Variance	Standard Deviation
1	325.47	18.04
2	65.63	8.10
3	26.01	5.10
4	29.79	5.46
5	54.09	7.35

The results of a formal multivariate analysis of variance are shown in Table 8. Only "significance levels," e.g., the probability of observing an F

Table 8

Multivariate Analysis of Variance for SAIS Scales

Source	Multivariate Results df	Multivariate Results p-level	Univariate Significance Levels SAIS Scale 1	2	3	4	5
Between Classes							
Mean	5,8	--	--	--	--	--	--
Class Type Differences	5,8	0.25	0.36	0.03	0.27	0.94	0.13
Aides	5,8	0.70	0.46	0.69	0.30	0.29	0.84
Size	5,8	0.02	0.42	0.30	0.77	0.92	0.60
All Interactions	20,27	0.94	0.36	0.86	0.28	0.74	0.98
Within Classes							
Linear Time	5,8	0.002*	0.002*	0.008*	0.36	0.71	0.06
Linear Time x Class Type	5,8	0.25	0.002*	0.09	0.22	0.67	0.02
Linear Time x Aides	5,8	0.047*	0.11	0.60	0.15	0.06	0.16
Linear Time x Size	5,8	0.002*	0.06	0.38	0.04	0.02	0.008*
Linear Time x All (1) Interactions	20,27	0.08	0.05	0.11	0.054	0.15	0.47
Non-Linear Time	5,8	0.25	0.09	0.36	0.38	0.76	0.56
Non-Linear Time x Class Type	5,8	0.85	0.78	0.75	0.34	0.86	0.96
Non-Linear Time x Aides	5,8	0.007*	0.08	0.16	0.19	0.08	0.10
Non-Linear Time x Size	5,8	0.09	0.60	0.06	0.80	0.70	0.99
Non-Linear Time x (1) All Interactions	20,27	0.26	0.63	0.04	0.05	0.20	0.63

(1) In no case were interactions - when the 4 contrasts per variate were split and tested separately - found to be significant. The univariate df were 1 for hypothesis and 12 df for error except for the case of interactions where the hypothesis or numerator df was 4.

statistic as large or larger than the one observed when the null hypothesis is true, are recorded. To reject the null hypothesis that a particular difference was zero required, in this analysis, a significance level of 0.05 or lower for the multivariate test. In the case of the univariate tests, the rule adopted was to require a p-level of 0.01 or lower to reject the null hypothesis.

The results of this analysis indicated that:

(1) Between class variations were small as none of the F-ratios were found to be significant (Table 8).

(2) Overall change in the scores did occur over the two testing periods. This change was essentially linear with a positive slope. Since a higher score on all scales (with the exception of Scale 2) indicates more undesirable behavior, we can conclude that negative behavior on the average increased over the project with the greatest increase occurring between T_1 and T_2 (see Table 9). Table 9 reveals that all the scales except Scale 2

Table 9

Least Square Estimates of [illegible] Scale Means for the Time Effect and the Aide x Time Effect (Estimates for the Full Model)

[illegible] Scale	Factor Level	1	2	3
1	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]
2	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]
3	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]
4	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]
5	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]
Average of All Scales	Aide	[illegible]	[illegible]	[illegible]
	No Aide	[illegible]	[illegible]	[illegible]
	Average	[illegible]	[illegible]	[illegible]

To simplify presentation, a constant of 100 was added to all means since Scale 2 was negative.

operated in much the same way, i.e., a general increase with the sharpest increase occurring between T_1 and T_2. Scale 2, Attentive Involvement in Classroom Activities, operated in the opposite manner, for there was an increase in this positive attribute over time (shown as a score decrease in Table 9 because of previous multiplication by -1). Scale 2 apparently operates independent of the other four scales (see low intercorrelation in Table 6).

(3) The amount of change during the entire study period and the relative change in the first period as compared to the second varied with the presence or absence of aides. The former is reflected in the statistical significance of the "linear" time x aide interaction shown in Table 8 while the second is indicated by statistical significance of the non-linear time x aide interaction. The difference in rates of change for the two conditions was not statistically significant for any of the five scales. This suggests that this source of variation was complex and not merely a simple difference reflected on all five scales in the same way. Estimates of expected responses for each scale at each of the three times observations were taken are shown in Table 9. In general, the significance of the linear component appears to reflect the fact that the observed change over the two periods was greater for classes without an aide. However, classes with aides showed increases in both periods; those without aides increased only in the first period which contributed to the non-linear time interaction with aides.

Summarizing to this point, the data showed significant overall increase in undesirable behavior over the three semesters. The greatest increase occurred between T_1 (middle of school year) and T_2 (beginning of fall semester). Moreover, the aides apparently had a beneficial effect in reducing the increase in undesirable behavior overall but particularly at the start of the fall semester.

(4) The magnitude of "linear" change was associated with the size of the class. The differences in linear trend were fairly large for all scales except 2. However, the univariate difference was statistically significant for Scale 3 only (see Table 8). Estimates of the expected responses for the Size x Time Interaction are shown in Table 10. For scales 3, 4, and 5, the change over time for the standard size classes was small compared to the positive though moderate slope, which characterized the increased size classes. For scale 1 an overall increase (T_1 to T_3) scores occurred for both levels of the factor but the increase occurred most markedly and over all periods for the larger classes.

In sum, class size had a significant effect on the rate at which the undesirable behavior increased over the three semesters, with the increased class size having a much more negative effect than did the standard size classes. Moreover, the behavior most affected appeared to be the marked increase in Withdrawn Behavior (scale 3) which characterized the larger sized classes.

Table 10

Least Square Estimates of BASE Scale Means for the Class Size x Time Effect (Estimates for the Full Model)

BASE Scale	Factor Level	Time 1	Time 2	Time 3
1	Standard	137.37	163.06	163.70
	Decreased	137.83	166.25	169.76
2	Standard	57.03	57.73	53.86
	Decreased	60.82	58.47	59.87
3	Standard	119.36	118.49	119.11
	Decreased	117.71	118.40	119.42
4	Standard	119.40	119.98	119.49
	Decreased	119.27	120.33	120.43
5	Standard	128.41	127.43	128.11
	Decreased	119.21	129.83	131.82
Average (Weighted)	Standard	116.41	117.38	117.23
	Decreased	116.43	118.70	119.92

To clarify presentation, a constant of 100 was added to all means since Scale 2 was negative.

(3) While the multivariate test of the Class Type x Time Interaction was not significant, the univariate test for scale 1 was significant at the .01 level (actual p=.0011). Below are displayed the least square estimates of the means:

	Time 1	Time 2	Time 3
Primary Class	53.43	63.22	67.02
Intermediate Class	61.73	66.08	66.44

It is readily apparent that while the primary classes exhibited less Disruptive Behavior than the intermediate classes at the beginning of the project (February, 1967), by the end (June, 1968) their behavior was every bit as unruly as the intermediate pupils. Also, in the latter two semesters primary pupils continued to manifest increasing disruptive behavior while the intermediate pupils did not.

Film Scores

Three measures of class behavior were obtained from the film monitoring of the project. These were ambient noise level (NL), measured in db, activity level (AL), a general index of overall classroom activity at time of filming, and non-task-oriented behavior (NTO), an index of distractibility or lack of attention. These three measures were averaged for each class for each of the three semesters that comprised the length of the study. The major objective was to determine if these scores changed with time and if the differences in changes over time were related to the independent variables in the study or their interactions.

Preliminary Analysis

To simplify the analysis it was decided to determine whether the non-linear component of any time trend could be regarded as zero. The criterion for rejection of this null hypothesis was significance at the .15 level. The results are shown in Table 11. Only in the case of NTO scores was it necessary to reject the null hypothesis. Thus, the remainder of the analysis does not include the non-linear component for NL and AL but only the linear component, while in the case of NTO all components are studied. The results are shown in Table 12.

Table 11

Non-Linear Time Trend Analysis

Variable	Sum of Squares For Non-Linear Trend	df	Mean Square	df Error	Mean Error Square	F-Ratio	Significance Level
NL	162.76	8	20.34	13	25.62	0.79	p=.72
AL	591.08	8	73.88	13	67.83	1.09	p=.50
NTO	196.08	8	24.51	13	11.30	2.17	p=.13

NL = Noise Level   NTO = Non-Task-Oriented Behavior   AL = Activity Level

Between Class Analysis

Interactions

To determine if a further simplification could be achieved the first analysis tested the null hypothesis of no first order interactions among class characteristics (but not time). For all three of the scores it was not necessary to reject this. Thus, only main effects needed to be considered hypotheses.

Class Type

The null hypothesis was that primary and intermediate classes did not differ. This hypothesis could not be rejected (using the .05 level

Table 12

Analysis of Variance of Film Indices

a) Median Level (ML)	Sum of Squares	df	Mean Square	F-Ratio	Significance Level
Between Classes					
Mean	—	20	—	—	—
Class Type	263.23	1	263.23	10.42	p=.007
Aides	73.23	1	73.23	2.93	p=.110
Size	58.03	1	58.03	2.33	p=.150
Interactions	260.68	4	65.17	2.58	p=.230
Error (Classes Within Cells)	303.70	12	25.31	—	—
By Time (Linear)					
Time	703.82	1	703.82	77.54	p<.001
Interactions x Time	300.03	4	75.01	8.26	p=.002
Class Type x Time	76.36	1	76.36	8.41	p=.013
Aides x Time	0.22	1	0.22	0.02	p=.880
Size x Time	0.08	1	0.08	0.01	p=.920
Class Type x Aides x Time	7.11$^{(a)}$	1	7.11	0.78	p=.394
Class Type x Size x Time	30.37(a)	1	30.37	3.34	p=.092
Aides x Size x Time	143.03(a)	1	143.03	15.75	p=.002*
Class Type x Aides x Size x Time	3.37(a)	1	3.37	0.37	p=.560
Error (Class x Time Within Cells)	108.93	12	9.08	—	—

(a) Sums of Squares obtained after removing main effect sums of squares and in the order shown.

* All significance levels are nominal in that they would hold only if the particular effect had been singled out *a priori* to be tested in this manner. Actually, these F-ratios were computed only after the significance of the total interactions x time sum of squares had been determined. To be significant a particular test must be significant at the .05/4 = .0125 level. Only the aides x size x time interaction was significant.

Table 13 (con't)

b) Activity Level (AL)	Sum of Squares	df	Mean Square	F-Ratio	Significance Level
Between Classes	—	20	—	—	—
Mean	—	1	—	—	—
Class Types	17.73	1	17.73	0.12	p=.73
Ability	309.08	1	309.08	2.16	p=.17
Size	136.79	1	136.79	0.96	p=.35
Interactions	380.43	4	95.11	0.67	p=.50
Error (Classes Within Cells)	1715.29	12	142.94	—	—
By Time (Linear)					
Time	72.90	1	72.90	1.12	p=.26
Class Type x Time	179.49	1	179.49	1.58	p=.24
Ability x Time	99.43	1	99.43	2.06	p=.18
Size x Time	208.11	1	208.11	3.19	p=.10
Interactions x Time	75.06	4	18.76	0.38	p=.80
Error (Classes x Time Within Cells)	581.48	12	48.46	—	—

c) Non-Task-Oriented Behavior (NTOB)	Sum of Squares	df	Mean Square	F-Ratio	Significance Level
Between Classes	—	20	—	—	—
Mean	—	1	—	—	—
Class Types	5.87	1	5.87	0.17	p=.73
Ability	1.48	1	1.48	0.04	p=.77
Size	43.90	1	43.90	1.27	p=.28
Interactions	216.13	4	54.03	1.56	p=.23
Error (Classes Within Cells)	413.36	12	34.45	—	—
By Time (Linear and Non-Linear)					
Time	161.49	2	80.74	7.12	p=.0035
Linear	68.38	1	68.38	6.03	p=.03
Non-Linear	93.10	1	93.10	8.21	p=.01
Class Type x Time	42.36	2	21.18	1.87	p=.18
Ability x Time	3.04	2	1.52	0.13	p=.88
Size x Time	46.28	2	23.14	2.04	p=.15
Interactions	79.08	8	9.88	0.88	p=.55
Error (Class x Time Within Cell)	272.03	24	11.34	—	—

as the criterion) in the case of RTD and AL. The F-ratio for WL was significant at the 0.007 level (Table 12). Estimates of various differences are shown in Table 13a. The primary classes had higher means than the intermediate for WL, but the reverse occurred for RTD and AL, although the last two differences were not statistically reliable.

Table 13

Estimates of Differences of Film Indices

	Estimate			Standard Error		
Source of Difference	WL	AL	RTD	WL	AL	RTD
a) Mean	108.03	68.11	16.38	1.22	2.09	1.42
Primary vs. Intermediate	5.03*	-1.13	-1.22	2.44	3.79	2.06
Aides vs. No Aides	-0.36	-7.10	0.33	2.44	3.79	2.06
Standard vs. Increased	-1.31	-2.41	-1.07	2.44	3.79	2.06
Time						
Linear	-5.06*	-2.32	2.10*	0.73	1.40	0.82
Non-Linear	1.48	1.76	2.38*	1.23	1.99	0.81
b) Slope Differences						
Primary vs. Intermediate	4.13*	-1.49	1.09	1.46	3.37	1.63
Aides vs. No Aides	0.23	4.13	-0.21	1.46	3.37	1.63
Standard vs. Increased	0.13	-4.29	-1.05	1.46	3.37	1.63
c) Differences in Change of Slope						
Primary vs. Intermediate	-0.37	-3.32	-2.41	2.43	3.98	1.63
Aides vs. No Aides	-0.71	-0.93	-1.03	2.43	3.98	1.63
Standard vs. Increased	-2.39	-2.44	-3.09	2.43	3.98	1.63

*Null hypothesis that true value was zero has been rejected in the analysis.

Aides

The null hypothesis that the presence or absence of aides did not affect responses markedly could not be rejected for any of the dependent variables.

Size

The hypothesis of no difference between standard and increased classes could not be rejected. However, in all cases, the scores for the increased classes were higher.

Within Class Analysis: Interaction of Time with Other Factors

As a preliminary step, the hypothesis that higher order interactions with time had no effect was tested. This hypothesis could not be rejected for AL and STD (Table 12b and 12c), but could for SL (Table 12a). Thus, for AL and STD only main effects had to be considered while for SL the interactions were examined. In Table 12a it can be noted that the interactions found to be significant were removed before the main effects. This strategy was deliberate since the main effects were the major concern of the study. The strategy required two passes at the data but provided a more sensitive test of the hypothesis concerning the main effects. The non-linear time x class interaction was dropped from the error term in two analyses (Table 12a, c).

Time

In the case of SL and STD it was necessary to reject the null-hypothesis of no change over time. In the case of STD both the linear and non-linear time trend effects were significant using the S-method to evaluate the significance of both contrasts.

For SL and AL the slope of the time line was negative indicating a decrease in scores with time. The slope for STD was positive and the "quadratic" contrast was also positive indicating an increase in scores with time and a larger increase in the second period.

Class Type x Time

In only the case of SL was the difference in slope of the time line on the entire time curve between primary and intermediate classes significant. In this case the slope for primary classes was less steep (less negative) than for intermediate classes. However, SL scores for primary classes did decrease with time. For AL the difference was not significant but the slope for primary classes was more negative than for intermediate classes indicating that scores decreased more with time for the former classes.

The STD slope for primary classes was greater (more positive) but the increase in the slope of the total curve in the second period was not as great as for intermediate classes. However, the estimates for AL and STD could not be considered statistically reliable.

Aides x Time

In no case does the data permit rejection of the null hypothesis that changes over time were greater or less for classes with aides than for those without them. The differences observed suggest that decreases in SL and AL scores were smaller for classes with aides. For STD, the increases were smaller for classes with aides. However, none of these estimates could be regarded as different from zero to a statistically significant degree.

Size x Time

As in the previous case, no differences which are statistically significant could be found between standard and increased classes with respect to change in scores over time.

Class Size x Aides x Linear Time

As a preliminary test, the total of all three or more factor interactions involving the linear component of time were examined. The sums of squares were pooled and were computed as though these effects had been removed from the model first. The interactions were found to be significant (see Table 12a) for NL only. Then, the four interactions of individual characteristics with linear time were examined as though removed from the model after the higher order interactions to avoid biases caused by sums of squares which could be attributed to either the higher order interactions or the one factor interactions with time.

To examine the higher order interactions to determine which, if any, could be singled out as significant in particular without similar biases, the order or removal of effects was reversed and the higher order interaction sums of squares were obtained as though the group had been removed last. (Variations in order of removal among the four components of the higher order interactions was not attempted. However, evidence from examination of least square estimates of effects and standard errors suggested that the same conclusions would be obtained regardless of order). This examination indicates that the aides x size x linear time interaction in particular was significant but none of the others can be so regarded.

The estimate of the aides x size x linear time contrast was -11.19 with a standard error of 2.92. Least square estimates of the linear time slope parameter for NL for each of the four conditions is shown in the following table:

	AIDES	
	Aides	No Aides
Standard	-8.44	-3.10
Increased	-2.99	-8.84

The classical interaction pattern is clearly apparent and indicates that the decrease in NL was greater for the standard classes with aides and increased classes without aides relative to increased classes with aides or standard classes without aides. It is also apparent that the amount of decrease of NL scores did not vary with either of the two factors alone.

Summary

Significant changes over time did occur on two of the measures of behavior. There was a decrease in noise level (NL) and an increase in non-task-oriented behavior (NTOB) with a greater increase between T_2 and T_3 than between T_1 and T_2 for the latter scores. Primary classes had a higher ambient noise level than the intermediate classes and over the study period the decrease in NL was less for the primary classes.

While a clear interaction between aides and size of class over time existed for NL, it was virtually impossible of meaningful interpretation. The presence of aides appeared to enhance the reduction of noise in the standard size classes while the lack of aides appeared to enhance the noise reduction to about the same extent in the increased size classes.

Absentee Analysis

Absentee rates averaged for the classes at each of three times were analyzed for (1) sources associated with variations in the absentee rates averaged over the three periods, and (2) variations in the rates over time. Whether repeated measure analysis of variance or multivariate analysis of variance was used the results were the same: no differences were found to be statistically significant.

Table 14

Analysis of Variance for Absenteeism

Source	*df*	*F-ratio*	*p-level*
Between Classes	20		
Mean	1	--	--
Class Types	1	0.26	0.63
Aides vs. No Aides	1	0.81	0.39
Standard vs. Increased Size	1	3.71	0.08
Class Type x Aides	1	0.73	0.41
Class Type x Size	1	0.08	0.79
Aides x Size	1	0.19	0.67
Class Type x Aides x Size	1	0.03	0.86
Between Classes Within Cells	12	MS = 21.40	--
Within Classes	40		
Time	2	0.31	>.10
Time x Class Type	2	0.82	>.10
Time x Aides	2	0.91	>.10
Time x Size	2	0.008	>.10
Time x All Interactions	2	0.58	>.10
Time x Classes Within Cells	24	MS = 4.31	--

c. Language

Illinois Test of Psycholinguistic Abilities (ITPA)

Raw scores on the nine subtests[3], of the ITPA, were obtained at the beginning and at the end of the study period. Thus, it was possible to estimate only the linear time trend component of any change in time with these scores. The adjective "linear" must be used with caution as it was not possible to test the hypothesis that all higher order components of any time trend were zero. If any were non-zero, the estimate of the linear component was biased and included not only the linear component but also those parts of non-linear components which were correlated with it. The ITPA total score was handled in the analysis as the average across all subtest and is referred to as such.

Preliminary Analysis

The analysis could be considerably simplified if it could be established that there were no significant variations among the nine ITPA scores in terms of differences between time periods or between cells. If this was the case, the average of the nine ITPA scores rather than the nine individual scores could be used to study these differences and the sensitivity of the analysis would be enhanced.

The nine scores at each time were averaged across time to obtain nine average scores on the individual tests for the entire study period. These nine means were then averaged to obtain a single average ITPA total score for each cell. To study variations among the scores, eight linearly independent deviations of the individual average scores from the total average score were obtained. In a similar manner, the differences on each individual test between the two periods were averaged to obtain an average difference over time for each cell. Then eight linearly independent deviations were obtained to study variations among the tests in changes over time.

The hypothesis of no variations among individual tests in terms of differences among cells, or differences by time, or differences by cell by time might be tested by a repeated measures analysis of variance to test the source "between tests." Such an approach would rely heavily on the robustness of that analysis to departures from the assumptions of equality of variances and equality of covariances among pairs of sub-tests. Instead multivariate analyses of variance were used. With so many variables and so few degrees of freedom for error, this analysis was not particularly sensitive to variations among the sub-tests; yet a full analysis of variance for all nine sub-tests would also be limited in the same way.

---

[3] ITPA Subtests
(1) Auditory-Vocal Automatic Test
(2) Visual Decoding Test
(3) Motor Encoding Test
(4) Auditory-Vocal Association Test
(5) Visual-Motor Sequencing Test
(6) Vocal Encoding Test
(7) Auditory-Vocal Sequencing Test
(8) Visual-Motor Association Test
(9) Auditory Decoding Test

Table 13

Multivariate Analysis of Variations Among ITPA Sub-Tests

Source Between Classes	F-Ratio Multivariate Test	df	Significance Level
Mean Differences Sub-Tests	163.83	8,3	p<.001
All Interactions x Sub-Tests	0.37	32,20	p<.97
Class Type x Sub-Tests	8.80	8,3	p<.05
Aides x Sub-Tests	0.47	8,3	p<.83
Size x Sub-Tests	6.66	8,3	p<.08
"Within Classes" By Time			
Time x Sub-Tests	3.36	8,3	p<.18
All Interactions x Time x Sub-Tests	0.66	32,20	p<.88
Class Type x Time x Sub-Tests	1.48	8,3	p<.39
Aides x Time x Sub-Tests	0.83	8,3	p<.63
Size x Time x Sub-tests	0.66	8,3	p<.71

Effects removed in order shown. Reversals in order did not markedly affect the results.

The main source of variation was in mean score over all cells and both test periods among the nine sub-tests (see Table 13). This was not surprising as equality in the mean scores among the tests was not expected. However, differences among cells or between time periods or between cells by time period on the individual tests were not large enough to necessitate rejection of the hypothesis of no differences at the 0.05 level. However, the differences between standard and increased classes varied enough among the nine tests so that this source of variation could be considered statistically significant[3] (see Table 13). The nature of this variation is further considered in the discussion of class type effect on average ITPA score.

Analysis of Variance of ITPA Scores Averaged Over All Tests Between Cells

Class Type

The hypothesis of no difference between primary and intermediate classes on average ITPA scores was rejected. The average score for the intermediate classes were higher than the primary classes. This was the case for each of the nine sub-tests averaged over the two periods as well. The difference was smallest for sub-test one and largest for sub-tests seven and eight, the ones with the greatest deviation from the average. The .01 level of significance was required for hypothesis rejection in the case of the univariate effects. Table 16 contains the analysis of variance and Table 17 the estimate of the effects discussed in the remainder of this section.

---

[3] A multivariate analysis of variance using all nine ITPA sub-tests was also done to specifically analyze differences between the standard and increased classes and other differences. The results do not vary markedly from those reported.

Table 16

Analysis of Variance of Average ITPA Test Scores

Source Between Classes	Sums of Squares	df	Mean Square	F-Ratio	Significance Level
Mean	—	1	—	—	—
Class Type	936.98	1	936.98	24.75	p<.001
Aides	23.86	1	23.86	0.63	p>.43
Size	2.44	1	2.44	0.06	p>.80
Interactions	131.58	6	21.93	0.58	p>.75
Error (Within Cell)	454.08	12	37.84	—	—
"Within Classes" by Time					
Time	422.38	1	422.38	68.83	p>.001
Time x Class Type	29.51	1	29.51	4.81	p>.05
Time x Aide	6.40	1	6.40	1.04	p>.33
Time x Size	0.06	1	0.06	0.01	p>.90
Time x Interactions	1.22	6	0.20	0.03	p>.90
Error (Class x Time Within Cell)	73.66	12	6.14	—	—

Table 17

Estimates of Effects for Average ITPA Score

Score	Estimate	Standard Error
Mean	88.738	1.49
Differences		
Primary vs. Intermediate	-12.78*	2.98
Aide vs. No Aides	-3.14	2.98
Standard vs. Increased	-0.38	2.98
Time	4.33*	0.60
Differences in Slopes		
Primary vs. Intermediate	2.40*	1.20
Aides vs. No Aides	1.28	1.20
Standard vs. Increased	0.14	1.20

* Null hypothesis that the difference estimated was zero has been rejected.

Aides

The hypothesis of no difference between classes with aides and those without aides could not be rejected. The average score was higher for the latter and this was the case for all nine individual tests. However, none of these observed differences could be regarded as large enough relative to the error variation to be indicative of true differences.

Size

The null hypothesis of no difference between standard and increased classes could not be rejected. The average score was slightly higher for increased classes. However, this was not the case for all sub-tests. The sub-tests for which the reverse was the observed case were, in order of magnitude, 6, 7, and 3, while the greatest difference "in favor of" the increased classes occurred for test 2.

None of the differences on the individual ITPA tests was especially large nor would any be statistically significant if examined individually. The significant variation between tests noted in the preliminary analysis was the result of the reversals in trend noted.

Interactions

The null hypothesis of no differences associated with interactions of the characteristics of the classes could not be rejected.

Child Gains

Interactions of Cell Characteristics with Time

The hypothesis of no higher order interactions of class characteristics and time could not be rejected so that the remainder of the analysis can be focused on first order interactions and the main effect of time.

Time

The null hypothesis that scores at the end of the study period did not differ from those at the beginning of the study was rejected. The average score increased with time. Variations among the tests were small in this respect.

Time x Class Type

The null hypothesis that the change over time for primary and intermediate classes was the same was rejected. The scores for primary classes increased more during the study period than the intermediate classes. However, the increase was not sufficient to reverse the overall difference between the two class types.

Time x Aides

The hypothesis that the change over time was the same regardless of the presence or absence of aides could not be rejected. The estimated difference in slopes was positive, however, indicating that the observed change was greater for the classes with aides. However, this could not be regarded as a statistically reliable difference.

Time x Size

The hypothesis of no difference in change over time between standard and increased classes could not be rejected. The observed difference was quite small.

Conclusions

When the order in which these differences were analyzed was changed, the results did not change in any meaningful manner. Thus, the major conclusions were that the ITBS total scores increased during the study period. The increase was greater for primary classes, but primary classes had lower scores than intermediate classes at all times.

Linguistic Analyses

The stimulus-elicited language samples were analyzed to yield 19 indices which on the basis of previous research and a priori considerations would be influenced by the independent variables of the study. The derivation procedures and supporting studies are contained in Appendix C.

The original 19 indices were subjected to scale modifications so that all indices would have the same (about 1.0) variance and be in the same direction. The scale multiplier used are shown in Table 18. This was done

Table 18

Scale Changes in Linguistic Analysis

Index No.	Description	Scale Multiple
1	Words per minute	0.037
2	Mean words per narrative segment	0.7167
3	Ratio of syllables in excess to total syllables in transcript	11.0
4	Ratio of type to tokens	9.5
5	Ratio of acceptable tokens to total tokens	-43.0
6	Ratio of unacceptable phonological tokens to total tokens	29.0
7	Ratio of unacceptable morphological tokens to total tokens	115.0
8	Ratio of unacceptable usage tokens to total tokens	139.0
9	Ratio of unacceptable English word tokens to total tokens	39.0
10	Ratio of unresolved nouns to total nouns	9.0
11	Ratio of incorrect verb forms (including omission of "to be") to total finite verbs	4.5
12	Ratio of incorrect verb forms which are omissions of "to be" with main verb following to total finite verbs	7.5
13	Ratio of incorrect verb forms which are omissions of "to be" as main verb (no verb in sentence) to total finite verbs	7.95
14	Ratio of number of type 1 communication units to total number of tokens	17.0
15	Ratio of number of type 2 communication units to total number of tokens	59.0
16	Ratio of number of type 3 communication units to total number of tokens	11.55
17	Average number of syllables per token	27.00
18	Ratio of appropriately spoken complete intonation patterns to total number of complete intonation patterns	5.55
19	Ratio of appropriately spoken independent clauses to total number of independent clauses	

so that the relative contribution of each index to a sum or difference would not depend on its variance. That is, the "noisy" indices could "swamp" the others.

Composite Indices

Composite indices, four in number, were created by summing the scale modified indices. The composite indices were Verbal Output, Pauses, Intonation, and Acceptability. The original indices used to make up the four composite ones are shown in Table 19. The desirability of increases in scores is also indicated.

Table 19

Grouping of 19 Derived Indices and Desirability
of Increased Magnitude of Scores

Group	Index # From Table 18		Desirability Direction
Verbal Output	1		+
	2		+
	4		+
	17		+
Pauses	3		−
	10		−
Intonation Patterns	14		+
	15		+
	16		+
	18		+
	19		+
Acceptability Measures	5	(after scale change)	−
	6		−
	7		−
	8		−
	9		−
	11		−
	12		−
	13		−

+ Increases in score associated with desirable change
- Increases in score associated with undesirable change

The four composite indices were subjected to a further scale modification for the same reasons as before; however, because of varying correlations among the original indices making up the linear combinations used to obtain the new ones, the variances of the four indices differed somewhat as may be seen in Table 20.

Table 20

Means, Variances, and Standard Deviations of Composite Indices and Correlations Among Indices

Index	Mean	Variance	Standard Deviation	
Verbal Output	2.3	0.21	1.56	0.45
Pauses	-0.07	0.29	1.10	0.54
Intonation	3.15	0.03	0.91	0.17
Acceptability	7.29	0.17	1.26	0.41

df = 73

Correlation Matrix (Within Cells Averaged) Index

	Verbal Output	Pauses	Intonation	Acceptability
Verbal Output	1.00			
Pauses	0.05	1.00		
Intonation	-0.43	0.10	1.00	
Acceptability	0.36	-0.06	-0.30	1.00

df = 73

The final scale codification is shown in Table 21.

Table 21

Composition of Final Indices and Scale Modification on Each

Composite Index	Scale Multiple	Combined Indices
Verbal Output	0.25	1, 2, 4, 17
Pauses	-0.05	3, 10
Intonation	0.30	14, 15, 16, 18, 19
Acceptability	-0.125	5, 6, 7, 8, 9, 11, 12, 13

Multivariate Analysis of Variance

The four dependent variables, the four composite indices, were subjected to a multivariate analysis of variance to determine which of any sources of variation or differences might be regarded as real. The null hypothesis of difference in response associated with a difference in condition was rejected, for multivariate tests, if the test statistic attained a size such as to be associated with a probability less than or equal to 0.05 if no difference actually existed. For univariate tests the criterion was a probability of less than or equal to 0.05.

The results of the multivariate analysis are shown in Table 22. The main results were:

Table 22

Multivariate Analysis of Variance: Linguistic Indices

Source	Multivariate Test df	Multivariate Test p-level	Univariate Significance (p-level) Speech Index Verbal Output	Vagues	Interr.	Accept.
Between Classes						
Mean	4,69	--	--	--	--	--
Class Type Difference	4,69	0.09	0.09	0.42	0.54	0.01
Aides	4,69	0.26	0.36	0.15	0.30	0.32
Size	4,69	0.23	0.11	0.13	0.15	0.86
Class Type x Aides	4,69	0.56	0.36	0.87	0.40	0.70
Class Type x Size	4,69	0.001*	0.05	0.19	0.06	0.002*
Aides x Size	4,69	0.31	0.89	0.62	0.32	0.17
Class Type x Aides x Size	4,69	0.35	0.07	0.97	0.93	0.96
"Within" Classes						
Linear Time	4,69	0.001*	0.001*	0.005*	0.012	0.001*
Linear Time x Class Type	4,69	0.003*	0.007*	0.13	0.67	0.008*
Linear Time x Aides	4,69	0.08	0.14	0.03	0.57	0.61
Linear Time x Size	4,69	0.39	0.09	0.46	0.32	0.75
Linear Time x Class Type x Aides	4,69	0.50	0.48	0.63	0.71	0.31
Linear Time x Class Type x Size	4,69	0.30	0.21	1.00	0.41	0.85
Linear Time x Aides x Size	4,69	0.01*	0.56	0.10	0.67	0.002*
Linear Time x Class Type x Aide x Size	4,69	0.08	0.16	0.97	0.97	0.007*
Non-Linear Time	4,69	0.001*	0.001*	0.71	0.03	0.008*
Non-Linear Time x Class Type	4,69	0.07	0.36	0.39	0.003*	0.48
Non-Linear Time x Aides	4,69	0.81	0.29	0.49	0.73	0.96
Non-Linear Time x Size	4,69	0.42	0.83	0.82	0.11	0.97
Non-Linear Time x All Interactions	8,211	0.42	0.18	0.17	0.83	0.38

*Criterion for Significance
Multivariate test: An observed p-level of 0.05 or less
Univariate test: An observed p-level of 0.015 or less

Degrees of Freedom for Univariate tests: 1 for hypothesis and 72 for error except for non-linear time x all other interactions in which case the hypothesis df was 3.

(1) Few "between" class differences were found. (These were differences in the mean values of responses averaged over the three test occasions.) The only difference large enough to be considered statistically significant was that associated with the class type x size interaction. In the absence of other significant differences, one might conclude that: (a) the responses for primary students in standard sized classes and intermediate students in increased sized classes were little different, but differed from those in classes which were primary of increased size and intermediate of standard size, which do not appear to differ substantially. For Verbal Output the estimated typical response was higher for the first pair while for Mazes, Intonation, and Acceptability the typical response was lower for the first pair of conditions. These estimates are shown in Table 23 with the standard errors. In all cases the estimate was large relative to the standard error; this implied that if the true value estimated was zero, then the observed estimates were substantial deviations. However, only in the case of Acceptability was it necessary to reject the null hypothesis at less than the .015 level of confidence (z = 2.75, p<.006).

Table 23

Least Square Estimates of Class Type x Size
Differences and Standard Errors

Index	Estimate(a)	Stand. Error
Verbal Output	0.69	0.35
Mazes	-0.11	0.19
Intonation	-0.28	0.13
Acceptability	-0.88*	0.32

(a) The metric was proportional to but not the same as for Table 21. The estimate was the estimate of one-half the difference between primary standard classes plus secondary increased size classes vs. primary increased classes and secondary standard size classes.

* The univariate test of the null hypothesis that the true value estimated was zero was rejected in the analysis of variance.

(2) The null hypothesis of no change over time was rejected both for the "linear" and "quadratic" components of time trend (Table 22). It was not possible therefore to consider the change as only linear or quadratic. Since only three points in time were observed, it was not possible to rule out the existence of higher order components; hence, the change was termed simply "change" even though the component which would be termed "linear" may be singled out for examination as a convenience in the discussion.

The estimates of the "linear" component of change were positive for Verbal Output, Mazes, and Acceptability, but negative for Intonation. The estimates of the "quadratic" component were positive in all cases but Verbal Output. These are shown in Table 24. The values shown would imply: (a) Verbal output increased during the first period but decreased in

Table 24

Estimates of Time Trend Components

Index	"Linear" Estimate (i)	"Linear" Stand. Error (i)	"Quadratic" Estimate (ii)	"Quadratic" Stand. Error (ii)
Verbal Output	0.17*	0.047	-0.[illegible]	0.048
Pause	0.20*	0.049	0.[illegible]	0.101
Intonation	-0.25*	0.029	0.[illegible]	0.027
Acceptability	0.16*	0.031	0.[illegible]	0.032

* The univariate test of the hypothesis that the [illegible] value estimated was zero has been rejected at p<.01.

(i) Entries should be multiplied by $\sqrt{\frac{N}{2}}$ and

(ii) Entries should be multiplied by $\sqrt{6}$ to produce comparable entries.

the second period but were higher at the end of study than initially; (b) Pause responses increased in both periods, perhaps more in the second period; (c) Intonation responses were lower at the end of the second period than initially and decreased less during the second period or were lowest after the first period; and (d) Acceptability responses increased over the two periods and increased more during the second period. In fact, there was a decrease during the first period. In the case of Verbal Pause and Acceptability the estimated "quadratic" component was larger than the estimated "linear" component. Estimates of means representing these trends are shown as "averages" in Table 25.

Table 25

Least Square Estimates of Mean for the Class Type x Time Interaction

Composite Index	Class	Time 1	Time 2	Time 3
Verbal Output	Primary	7.93	8.35	8.36
	Intermediate	8.26	8.59	8.32
	Average	8.10	8.47	8.34
Pause	Primary	-0.97	-0.99	-0.83
	Intermediate	-1.26	-1.01	-0.81
	Average	-1.11	-1.00	-0.82
Intonation	Primary	3.43	3.25	3.35
	Intermediate	3.36	3.33	3.29
	Average	3.39	3.29	3.32
Acceptability	Primary	7.06	7.07	7.37
	Intermediate	7.44	7.28	7.52
	Average	7.25	7.18	7.45
Average	Primary	4.36	4.42	4.56
	Intermediate	4.45	4.55	4.58
	Average	4.41	4.49	4.57

(3) The linear time x class type interaction was significant (Table 22). Table 26 presents the least square estimates of the class type x time parameters for each of the four composite indices. Linear slope differences were found between the primary and intermediate classes for the overall effect and for Verbal Output and for Acceptability. The effect was the same, the primary classes exhibited a steeper slope. The linear effect for Intonation (though not significant) was negative while the quadratic effect was significant as the primary classes showed a decrease in the first period and an increase in the second being larger.

Table 26

Least Square Estimates of Class Type x Time Parameters of Linguistic Indices

Composite Index			
Verbal Output	Est.	0.243*	0.076
	S.E.	0.094	0.080
Pauses	Est.	-0.211	0.008
	S.E.	0.139	0.142
Intonation	Est.	-0.016	0.120*
	S.E.	0.038	0.046
Acceptability	Est.	0.169*	-0.043
	S.E.	0.082	0.105

* Hypothesis of no slope differences between groups rejected.

(4) The "linear" component of the aide x size interaction over time was significant (Table 23). Estimates of the differences are shown in Table 27.

Table 27

Estimates of Differences in "Linear" Component of Change Associated with Aides by Class Size Interaction

Index	Estimate(1)	Stand. Error
Verbal Output	0.116	0.187
Pauses	-0.440	0.277
Intonation	0.032	0.076
Acceptability	0.323*	0.123

(1) The estimates are of the difference in "linear" components between standard classes with aides and increased classes without aides vs. standard classes without aides and increased classes with aides.

*The hypothesis that the true value estimated was zero has been rejected.

In all cases but Acceptability the estimates were small relative to the standard errors of the estimates, however, that for Noise was greater than the standard error in absolute magnitude but it was not statistically significant. For Acceptability the estimate was large and positive which implies that the estimated change for standard sized classes with aides or increased sized classes without aides was greater than for standard sized classes without aides or increased classes with aides. Estimated mean values for each time are shown in Table 28. The "linear" component was the difference between time 3 and time 1. The values shown for Acceptability indeed illustrate the effect just described.

Table 28

Estimates of Response by Index by Time
in Presence or Absence of Aides
and for Standard and Increased Size Classes

Index	Condition		Time 1	Time 2	Time 3
Verbal Output	A	S	7.95	8.39	8.44
	A	I	8.28	8.58	8.44
	NA	S	7.96	8.40	8.17
	NA	I	8.20	8.51	8.37
Noise	A	S	-0.87	-1.09	-0.90
	A	I	-0.90	-0.90	-0.73
	NA	S	-1.36	-1.34	-0.82
	NA	I	-1.08	-0.59	-0.80
Interaction	A	S	3.43	3.36	3.32
	A	I	3.41	3.37	3.33
	NA	S	3.42	3.34	3.33
	NA	I	3.31	3.29	3.29
Acceptability	A	S	7.11	7.17	7.30
	A	I	7.45	7.37	7.30
	NA	S	7.38	7.16	7.36
	NA	I	7.05	7.09	7.42

A = Aide
NA = No Aide

S = Standard size
I = Increased size

(3) Another interaction of interest was the class type x aides x size "linear" trend component for Acceptability (Table 27). The estimate of this effect was 0.487 with a standard error of 0.244. This indicated that the difference described in the previous paragraph was greater for primary classes than for intermediate ones. In fact, the linear component for Acceptability was highest for primary classes with aides of standard size and primary classes without aides of increased size. The same difference was positive for intermediate classes, but was much smaller (see Table 28).

Table 19

Estimates of the Linear Component of Time Trend for Acceptability by Cell

Cell	Class Type	Aides	Size	Estimate of Linear Component for Acceptability (X)	Sign for Contrast (XX)
1	P	A	S	0.45	+
2	P	A	I	-0.06	-
3	P	NA	S	0.08	-
4	P	NA	I	0.42	+
5	I	A	S	0.08	-
6	I	A	I	0.13	+
7	I	NA	S	-0.11	+
8	I	NA	I	0.13	-

Designations: Class Type : P = Primary, I = Intermediate
Aides : A = Aide Present, NA = No Aide
Size : S = Standard, I = Increased

(X) Standard Error of estimate in each cell was approximately 0.093.
(XX) Each estimate was given this sign for determination of the contrast which estimates the class type x aides x size interaction.

Summary of Analyses of Dependent Measures

The foregoing presentation dealt with the analysis of the data using the dependent measures as the basis for categorizing the various analyses. In this section the data are summarized using the independent variables as the basis for discussing the results. The significance of between cell effects are not considered here, though they are included in the foregoing sections. The reason is simple--within the context of this study they were of very limited interest as the major hypotheses were concerned with the treatment effects over time.

A. Time

In all analyses the linear and non-linear components of time and the effects x time interactions were examined. The linear analysis examined the slope of line between T_1 and T_3 while the non-linear analysis examined differences in slope between T_1-T_2 and T_2- T_3.

(1) Academic

Raw scores in all three subtests of the WRAT - reading, spelling, and arithmetic - increased significantly over the three semesters. For the arithmetic scores the increase was greatest during the second time period (T_2-T_3).

(2) Behavior Indices

(a) [illegible]

A composite measure of undesirable behavior (the average of the five [illegible] scales) increased over the three semesters of the project. However, when the five scales were considered individually, a significant increase in Disruptive Behavior over time was noted. A significant increase in Attentive Involvement in Classroom Activities was also noted. The latter, although an increase in desirable behavior, was not of sufficient magnitude to reverse the increase in undesirable behavior of the other four scales.

(b) Film Indices

Noise Level (ambient) decreased over the three semesters of the study while Non-Task Oriented behavior increased.

(3) Language

(a) ITPA

The average of the nine subtests of the ITPA (equivalent to the total language score in ITPA manual) increased over the course of the study.

(b) Linguistic Indices

The composite linguistic score (based on all 19 indices) when all indices were oriented in a "desirable" direction, increased significantly over time. The 19 indices were categorized into four composite indices (see Tables 18 and 19). Of these four, Verbal Output, Mazes, and Acceptability increased significantly over time while Intonation did not change significantly.

B. Class Type (Primary vs. Intermediate)

(1) Academic

The slope of the line over time for all three subtests of the WRAT was greater for the primary than for the intermediate classes.

(2) Behavior Indices

(a) BAPS

The class effect over time was not significant for the composite BAPS scores. However, Disruptive Behavior increased much more rapidly over the course of the study for the primary classes.

(b) Film Indices

Primary classes did not decrease in Noise Level to the extent that the intermediate classes did.

(3) Language

(a) ITPA

The increase in total language score was greater for the primary classes.

(b) Linguistic Indices

The overall increase in linguistic performance was greater for the primary classes. Individual composite indices whose slope was significantly greater for the primary than the intermediate classes were Verbal Output and Acceptability.

C. Aides (Aide in class vs. no aide in class)

(1) Academic

Aides had no effect.

(2) Behavior

(a) [illegible]

The presence of aides in the classes reduced the rate of increase of negative behavior over the three semesters as indexed by the average score of all five [illegible] scales.

(b) Film Indices

Aides had no effect.

(3) Language

(a) ITPA

Aides had no effect.

(b) Linguistic Indices

Aides alone had no effect.

9. Size of Class (Standard vs. Increased [standard + 20%])

(1) Academic

Arithmetic scores increased more over the three semesters for the standard sized classes.

(2) Behavior

(a) [illegible]

The overall increase in negative behavior was greatest for the increased class size with the individual scale of Withdrawn Behavior being most affected.

(b) Film Indices

Size alone had no effect.

(3) Language

(a) ITPA

Size alone had no effect.

(b) Linguistic Indices

Size alone had no effect.

E. Interactions Over Time

Only those interactions are noted which had a significant effect on the particular dependent variable. All the possible interactions with time were: class type x aides, class type x size, aides x size, and class type x aides x size.

(1) Class Type x Aides

Noise Level

Presence of aides in the standard sized classes resulted in a marked decrease in noise level compared to the absence of aides. On the other hand, *absence* of aides compared with aides in the increased classes resulted in a marked decrease in noise level.

(2) Class Type x Aides x Size

The linguistic index of acceptability was significantly affected by this interaction. The complexity of the interaction defied a clear-cut interpretation except to note that the responsivity to the aide and size factors was most pronounced in the primary classes.

2. Automatic Interaction Detector (AID) Analysis

In order to exhaust the information available in the data, it was subjected to an AID analysis using the program recently made available to the University of Wisconsin Computing Center's Univac 1108. The program abstract indicates:

> This program is useful in studying the interrelationships among a set of up to 37 variables. Regarding one of the variables as a dependent variable, the analysis employs a nonsymmetric branching process, based on variance analysis techniques, to subdivide the sample into a series of subgroups which maximize one's ability to predict values of the dependent variable. Linearity and additivity assumptions inherent in conventional multiple regression techniques are not required. ([illegible], 1970)

Appendix E contains a more detailed discussion of this type of multivariate analysis.

The analysis requires that one of the up to 37 variables be the dependent variable which can be either categorical or continuous. The change scores of the three WRAT subtests over the three semesters were used as dependent measures in three separate analyses. The predictor variables must be categorical in nature with a range from 0 to 63 categories. The predictor variables and their categorical transformation (where required) were as follows:

(1) Class - Primary 1
Intermediate 2

(2) Aides - Aides 1
No Aides 2

(3) Size - Standard 1
Increased 2

(4) IQ - 43-60 1
61-70 2
above 70 3

(5) MA - 50-70 1
71-90 2
91-110 3
above 110 4

(6) Teacher Sex - Male 1
Female 2

(7) Race - White 1
Non-White 2

(8) Pupil Sex - Male 1
Female 2

The following constraints were built into the analyses:

(1) .05 = proportion of the total sum of squares that must be contained in the i^{th} group if that group was to be eligible for splitting.

(2) The best split on the i^{th} candidate group must have reduced the unexplained sum of squares by the proportion .05 of the total sum of squares or that group did not become a candidate group again.

(3) 15 = the maximum number of final groups into which the input data could be split.

(4) 25 = minimum number of observations that must be contained in the i^{th} group if it was to be a candidate for splitting.

The priorities for imposing the constraints occurred in the order indicated. The N in the analysis was 171.

In summary, this analysis attempted to explain the characteristics of those groupings of pupils which accounted for the maximum amount of variance on the change scores of the three WRAT subtests. The characteristics were of necessity restricted to the eight predictor variables indicated above.

Reading Gain Scores

Using the above constraints, no splits from the original or total group were possible.

Spelling Gain Scores

No splits from total group were obtained.

Arithmetic Gain Scores

Splits were obtained on the arithmetic scores with the following mutually exclusive groups being formed.

Group(1)	Characteristics	N	Mean in Mos. Gain	SD
6	Standard size, primary	36	10.7	.07
5	Standard size, intermediate	43	7.4	.05
3	Increased size	92	6.3	.04
Total		171	7.5	.05

(1) The group number refers to the sequence in which the groups were generated by the computer program.

A one-way analysis of variance yielded the following:

Source	df	SS	MS	F
Between	2	492.6	246.2	11.0
Within	168	3768.1	22.4	

The obtained F was significant at better than the .005 level of confidence. The obtained grouping accounted for 11.6% of the variance.

Six of the eight predictor variables did not enter into the group characterizations since at each step one of the two variables used proved more useful in explaining the variance remaining in the particular group. In conclusion, prediction of arithmetic gain scores on the basis of membership in the three groups provided a significant reduction in the error.

Since the constraints in the analysis can be modified, it was considered informative to vary constraint #2 from .05 to .025. It was assumed that this reduction would maintain a certain robustness in the group characteristics yet at the same time allow for some notion of the characteristics of groupings of the pupils which might account for a significant portion of the variance of the reading and spelling gain scores in addition to the arithmetic scores.

Reading Gain Scores

With the smaller proportion of .025 for constraint #2, the following mutually exclusive groups evolved from the analysis of the reading scores:

Group	Characteristics	N	Mean in Mos. Gain	SD
8	IQ 45-60 mos., no aides, standard size	5	16.4	.11
9	IQ 61 mos.+, no aides, standard size, MA 50-90 mos.	20	7.3	.06
10	Aides, MA 91 mos.+	46	5.0	.04
5	Increased size, MA 50-90 mos.	48	4.8	.04
7	Aides, standard size, MA 50-90 mos.	20	4.6	.03
11	No aides, MA 91+	32	2.3	.03
Total		171	7.3	.05

The analysis of variance source table was:

Source	df	SS	MS	F
Between	5	752.0	150.4	7.40
Within	163	3309.9	20.3	

The obtained F of 7.40 was significant at better than the .005 level and the six groups accounted for 18.5% of the variance of the reading gain scores. Only IQ, MA, aides, and size entered into the characterization of these groups with class, teacher sex, pupil sex, and race not considered.

Group N's of less than 25, the built-in constraint, indicated the data do not conform to a unimodal distribution and that the characterization of these groups was probably quite unstable and therefore difficult to replicate. Indeed, it is difficult to understand why the groups that made the highest gains were characterized by the lowest IQ, while the group that had the least gain was characterized by the highest MA. A discussion of the utility of the MA and IQ as they relate to academic achievement is contained in Appendix A-5.

Spelling Gain Scores

The characteristics of the four groups which evolved from the analysis of the spelling scores was:

Group	Characteristics	N	Mean in Mos. Gain	SD
2	Aides	98	4.6	.04
4	Primary, aides	34	4.1	.04
6	IQ 61-70, intermediate, no aides	17	3.7	.02
7	IQ 45-60 and 91-110, intermediate, no aides	22	0.8	.02
Total		171	3.9	.03

The one-way analysis of variance yielded the following:

Source	df	SS	MS	F
Between	3	252.3	84.1	5.33
Within	167	2633.8	15.8	

The obtained F of 5.33 was significant at better than the .01 level and the four groups accounted for 8.7% of the variance on the gain scores. On looking at the groups, it would appear that having aides as opposed to not having aides in the class enhanced the spelling gains of the pupils.

Arithmetic Gain Scores

Within the constraints of this second analysis four groups resulted using the arithmetic scores.

Group	Characteristics	N	Mean in Mos. Gain	SD
4	Primary, standard size	34	10.7	.05
6	Aides, primary, standard size	27	8.7	.05
3	Increased size	92	6.3	.04
7	No aides, intermediate class, standard size	18	5.6	.04
Total		171	7.5	.05

Analysis of variance yielded:

Source	df	SS	MS	F
Between	3	607.2	202.4	9.25
Within	167	3653.3	21.9	

The obtained F of 9.25 was statistically significant (p<.001). With constraint #2 at .05, 11.7% of the variance was accounted for, while in this analysis with the constraint at .025, 14.3% of the variance was accounted for. The only difference between the two analyses was that group #5 characterized primarily by standard size, intermediate class split on aides and class (see groups 6 and 7 above).

Summary

When constraint #2 was equal to .05 the AID analysis results were similiar to those obtained in a formal multivariate analysis (summary of WRAT analysis on page 79). That is, EMR classes characterized as primary and of standard size obtained the greatest gain in arithmetic achievement. Differentiating characteristics were not evident for reading or spelling.

When constraint #2 was reduced to 0.25, the primary level, standard class size continued to characterize the group that gained most in arithmetic. The reading gains were not interpretable. In terms of spelling gains, groups characterized by being in classes with aides at the primary level made the greatest gains.

3. Cost Analysis

One of the original objectives of the study was to conduct a comparative cost analysis of classroom units with increased educable enrollment utilizing teacher aides versus classroom units with traditional enrollments and no aides. Of interest also was a cost efficiency analysis of the extreme case of increased class size without an aide versus the standard size with an aide. This section presents the results of analyses which attempted to answer the above questions.

Two approaches to the cost analysis were considered. In both instances certain assumptions had to be made and, if one could live with these assumptions, then the analysis made sense. The first approach considered was that of what can be called the "service" model. This model assumes that if a child can be brought into contact, spend most of his time with certain professional and paraprofessional personnel, all appropriately certified, then all sorts of positive things will start happening to him. An effective cost analysis under these circumstances should demonstrate that additional children can be "serviced" without an undue increase in the cost per child. The concern of the present project, under this model, was to show that the cost per pupil in an increased size class with an aide was not significantly greater than the standard sized EMR class without an aide.

The raw data for this and succeeding analysis is contained in Table 30. Summary data for the above concern was:

No Aides		Aides	
Standard Size	Increased Size	Standard Size	Increased Size
$912	$655	$1,132	$883

As the data shows, it costs about the same to serve a pupil one year in a standard size class without an aide ($912) as it does in a class increased in size, using an aide ($883). Indeed, the differential($29 per pupil) which did exist favored the latter type of class.

An unknown here is the value which resides in providing special education services to EMR children. If this value is large, then the dollar signs do not adequately reflect the worth of being taken off the EMR waiting list.

The second approach to the cost analysis was to use what might be called a "commerce" model. Here it is assumed that one knows what he wants to attain or produce and has limited resources to do so. Thus, the concern is to attain or produce a given unit of what is desired at the least possible cost. In the present study, the assumption

Table 30

Raw Data for Cost Analysis Based on the Two Semesters in the 1967-68 School Year

		No Aides				**Aides**			
		One Year Cost Per Pupil	Class Ave. One Year Grade Level Achievement (mos.)	Class Ave. Cost/Year Achievement	Ave. Size	One Year Cost Per Pupil	Class Ave. One Year Grade Level Achievement (mos.)	Class Ave. Cost/Year Achievement	Ave. Size
Primary	Standard Size	$1,075*	7.0	$1,536	10	$ 999	2.0	$4,993	11.5
		923	6.0	1,538	11.5	1,310	4.0	3,275	10.5
		934	4.7	1,987	11.5				
	Increased Size	$ 753	3.0	$2,510	16.5	$ 841	3.0	$2,802	13.5
						995	3.7	2,690	17.0
						1,293	3.0	4,308	13.0
Intermediate	Standard Size	$ 969	2.3	$4,215	14.5	$1,439	2.3	$6,257	10.5
		657	4.3	1,528	15.0	840	3.3	2,545	15.5
						1,199	5.7	2,103	12.5
						1,006	2.0	5,030	13.5
	Increased Size	$ 611	3.7	$1,651	17.5	$ 677	3.0	$2,257	21.5
		601	2.3	2,613	22.0	713	4.0	1,783	19.5
						782	4.3	1,819	20.5

---

*Based on teacher salary, aide salary (when applicable), administrative costs, transportation, books, instructive equipment, attendance, operational, maintenance and fixed costs, and school lunch. Exclusive of salaries the average cost per pupil was $192.

was made that academic achievement was trying to be maximized and a given unit of this achievement should be obtained at the lowest possible cost.

The procedure in obtaining a cost-achievement index for the 1967-68 school year was as follows: for a given class, teacher and aide (where applicable) salaries were obtained and added to the cost of administration, transportation, books, instructional equipment, attendance, operations, maintenance and fixed costs. The number of pupils in each class was averaged for T_1 (September, 1967) and T_3 (June, 1968). The average cost per pupil in each of the 20 classes in the project was then calculated. These figures are shown in Table 30.

The gain in achievement on the three WRAT subtest was calculated for the 1967-68 school year. The average gains on the three subtests were obtained. These figures are shown in the cells of the design in Table 30.

A meaningful cost-achievement index should be calculated on a common base. The common base selected was the academic school year of ten months. That is, using our data, the question was, "How much does it cost on the average in each class for an EMR pupil to achieve ten months?" The formula used was:

$$\frac{\text{Actual average achievement in 10 months/class}}{\text{Average cost per pupil in 10 months}} = \frac{\text{10 months of achievement}}{\text{Cost to achieve 10 months' growth (an unknown)}}$$

or:

$$\text{Cost to achieve 10 months' growth} = \frac{\text{10 (Average cost per pupil in 10 months)}}{\text{Actual average achievement during 1967-68 school year}}$$

The formula was calculated for each of the 20 classes and the results are displayed in the cells of the design in Table 30. The cost to move a pupil 10 months in achievement varied considerably from $1,110 to $6,757 (see Table 30). These figures appear substantial considering that in the regular classes in Wisconsin (using the same cost items) the cost per pupil was $624 in 1967-68. This comparison is meaningful if the not unreasonable assumption can be made that this average regular class pupil gains 10 months in achievement in one school year.

The cost achievement index developed here could be considered as simply another dependent measure and analyzed within the general research design of the study. However, the effects of certain contrasts were sought so this analysis was restricted to answering certain specific questions.

The treatment effects of interest were:

(A) Increased size, without aide

(B) Standard size, without aide

(C) Increased size, with aide

(D) Standard size, with aide

The questions of interest were:

(1) Is the average cost of an EMR pupil achieving an academic year greater in a standard sized class without an aide than in a class of increased size with an aide? The null hypothesis was $\overline{X}_B - \overline{X}_C = 0$.

(2) Is the average cost of an EMR pupil achieving an academic year greater in a class increased in size without an aide than in a class of standard size with an aide? The null hypothesis was $\overline{X}_A - \overline{X}_D = 0$.

Where: $\overline{X}_A = \$2,230$ $\overline{X}_B = \$2,610$

$\overline{X}_C = \$2,161$ $\overline{X}_D = \$4,033$

Table 31 shows a source table for the analysis of the data.

Table 31

Source Table for Analysis of Variance of
Cost-Achievement Data

Source	SS	df	MS	F
Aide	4,562,835.0	1	4,562,835.0	3.19
Size	2,110,875.2	1	2,110,875.2	1.47
Aide x Size	2,369,941.3	1	2,369,941.3	1.66
Within	22,821,723.3	16	1,426,357.7	

None of the effects as analyzed here were significant at the .05 level (F = 4.49, df = 1,16). An F-ratio for the planned contrasts used:

$$F = \frac{(\overline{X}_C - \overline{X}_B)^2}{\text{Within error } \left(\frac{1}{N_C} + \frac{1}{N_B}\right)}$$

The obtained F of 0.38 did not allow us to reject the null hypothesis of no differences between $\overline{X}_C$ and $\overline{X}_B$. The obtained F of 4.41 falling below the required F of 3.39 did not allow us to reject the null hypothesis of no difference between $\overline{X}_A$ and $\overline{X}_D$. However, the obtained F of 7.84 ($p<.05$) in contrasting $\bar{X}_D$ with $\bar{X}_A + \bar{X}_B + \bar{X}_C$ suggested that it was significantly more expensive to attain one year academic achievement in standard size classes with no aide than it was in the other three combinations of class and aide.

Table 32 shows the two-way tables of the means of the cost achievement data for the three factors in the design class, aides, and size. These data have heuristic more than any other value since statistically significant differences were not obtained (see Table 31). Thus, the average cost for one year achievement across all classes was $2,272. The least expensive combination was no aides in the primary classes ($1,392) while the most expensive was aides in the standard classes ($4,033). Since these are only two-way tables, the third factor not shown in each table has been collapsed.

Table 32

Two-Way Tables of Means and Frequencies of Cost-Achievement Data for the Three Factors of Class, Aides and Size The Means are Based on Class Averages of the Dollar Cost of an ESE Pupil Achieving One Academic Year (10 Months)

	Class		
a) Aides	Primary	Intermediate	Across Class
Aides	$3614 (5)	$3113 (7)	$3322 (12)
No Aides	1992 (4)	2302 (4)	2197 (8)
Across Aides	2867 (9)	2891 (11)	2872 (20)

	Class		
b) Size	Primary	Intermediate	Across Class
Standard	$2166 (5)	$3612 (6)	$3182 (11)
Increased	3077 (4)	2023 (5)	2683 (9)
Across Size	2867 (9)	2891 (11)	2872 (20)

	Aides		
c) Size	Standard	Increased	Across Aides
Aides	$4513 (6)	$2161 (5)	$3182 (11)
No Aides	2633 (6)	2258 (3)	2683 (9)
Across Size	3322 (12)	2197 (8)	2872 (20)

B. Qualitative

The purpose of this section is to provide an overview of the teacher-aide dyad program in action, to report teacher and teacher aide subjective impressions, and to generate some working hypotheses regarding those features of the dyad which appear to be essential in an optimum teacher-aide relationship. Other sections of the final report document the more objective findings generated from an analysis of the resulting data. This section summarizes the impressions emanating from a consideration of teacher and aide responses to an interview questionnaire administered at the conclusion of the demonstration project. The sound tracks of filmed interviews were also analyzed. These teacher and aide reactions to the project and to the teacher-aide interactions were captured on the film Teacher Aides, produced and disseminated by the administering agency. The film is available for further study by the interested reader.

The research investigators anticipate that the information gained from these narrative reports and empirical impressions will complement the study findings and will be useful for administrative decision making. Information pertinent to the relevant motivational and interpersonal values held by generally successful team matchings can assist leadership personnel in modifying current administrative practices and decision-making processes as these relate to the use of paraprofessionals in special education.

When the project was initiated, paraprofessionals were a rare breed in special education with the exception of programs for the deaf. Wisconsin school districts were employing approximately a dozen teacher aides in school year 1966-67. This project through its dissemination and inservice components has had a statewide impact in demonstrating the feasibility of utilizing teacher aides in special education in general. Its influence is readily reflected in the data for paraprofessionals employed in special education in school year 1969-70 where the number has grown to 189 in three brief years. Careful analysis of the empirical impressions and reactions of successful teacher-aide dyads and teams by local leadership personnel should help to reduce the error in teacher aide assignments and function designations.

Each of the (1) general observations is followed by a (2) basic premise summarizing the hypothesis generated by subjective observations of teachers and aides interviewed regarding the project. These basic premises are, in turn, followed by (3) a series of direct quotes of administrators, teachers, and aides surveyed during and at the close of the project. Each quote has been included in a cluster of positive and negative reactions to the project and to the teacher aide concept in general. The quotation is preceded by a designation of teacher (T), administrator (A), or teacher aide (TA). Thus, the reader is afforded impressions and reactions from the real world of the teacher-aide dyad while the basic premises are documented by actual statements of project participants.

General Observations

In its original conception this research investigation was concerned with the manpower shortage in special education, with the need for new innovative approaches in educational arrangements utilizing paraprofessionals, and with the feasibility of increasing traditional primary and intermediate EMR class enrollments without impairing, and possibly, actually improving the learning efficiency of pupils enrolled in the participating units. The review of the literature on teacher-aide programs (see Chapter II) has further clarified the purposes of teacher-aide programs and has helped to conceptualize their basic intents. The literature suggests that the purposes of a teacher-aide program can be collated under two relatively distinct categories, each of which have administrative and research significance. These include:

(1) purposes which are primarily teacher-pupil oriented

(2) purposes which emphasize needs of a given subpopulation from which teacher aides are recruited and trained

This project was clearly conceptualized as being teacher-pupil oriented, one in which the professionalization of teaching and the individualization of learning was paramount. The provision of employment and new career opportunities involving untapped human resources was only a secondary concern of the project.

More specifically, teacher-aide programs with teacher-pupil oriented purposes enable the classroom teacher to devote more time to professional activities. Four primary goals can be consistently identified:

(1) To enable the teacher to devote more of her time to her teaching responsibilities;

(2) To enable more individualized instruction;

(3) To relieve the teacher of clerical duties; and

(4) To enrich child-adult relations, i.e., a second adult gives the child an alternative person to relate to, to work with, and to serve as a model.

Blessing (1967), Green (1966), and Cruickshank and Haring (1957) have all emphasized these goals as a basis for utilizing teacher aides in special classrooms. Blessing (1967, p. 107), for example, emphasized the notion that ". . . an aide in a classroom can relieve the teacher of a tremendous amount of work and facilitate the professional task of the instructor."

An additional purpose for introducing a paraprofessional into a special education program has been the possibility of increasing the number of students per teacher, thus reducing the critical manpower problem in this area. Green (1966) and Blessing (1967) have both made references to this potential value of teacher aides.

How well these goals were achieved in the current project is partially clarified in the quantitative results and discussion sections of this report. The admittedly more biased reactions of the participating teachers and paraprofessionals to some of these stipulated goals will be reported in subsequent portions of this section.

However, it should be noted that exposure of paraprofessionals to special education learning environments over the period of the project did, in fact, lead to the alternate purpose of providing employment and new career opportunities for a number of the teacher aides in the project. Actually the use of aides on the basis of their ability to satisfy particular situational needs has opened up a valuable sociological concept, i.e., the employment of economically and/or disadvantaged persons in school tasks appropriate to their competence. While the primary reason for employing such personnel was to enhance learning situations, it soon became readily apparent that their employment as an aide was accompanied by a dignity and integrity that injected a positive belief in upward mobility. Initially these teacher aides were employed at the hourly rate of $2.05. Upon completion of the project, all twelve remaining aides were absorbed within the Milwaukee Public Schools' special education mainstream as continuing employees under local and state support. As of January 1, 1970, these aides were reportedly receiving $3.16 per hour. Currently the maximum hourly rate they can earn after one year's experience as an aide is $3.36.

PREMISE I: THE UTILIZATION OF TEACHER AIDES IN SPECIAL EDUCATION PROVIDES EMPLOYMENT AND NEW CAREER LADDER OPPORTUNITIES, PARTICULARLY AMONG THE ECONOMICALLY AND/OR EDUCATIONALLY DISADVANTAGED. TEACHER AIDE EMPLOYMENT IS ACCOMPANIED BY A DIGNITY AND POTENTIAL FOR UPWARD MOBILITY.

Positives

T -- ". . . and let's face it, we are not that good in all of these areas (music, art, physical education). My inadequacies I am willing to face up to and I am crying for help. I need help in Phy.Ed., I need help in music, I need help in art . . ."

A -- "The paths of upward mobility of the aide should be explored, i.e., what is the end road for her? Is it a superaide, or should the way be deliberately opened to allow the aide to become part of the teaching profession on par with teachers?"

A -- ". . . it is recognized that teacher aide employment is accompanied by a dignity and integrity that injects a solid hope of upward mobility."

TA -- "At the same time it requires the type of personality of someone who is willing to take directions."

Negatives

A -- "The opinion of salary seems to be a point of concern. This should be considered from both an absolute point of view and also a relative point of view with respect to the salary differential between the teacher and the aide."

A -- ". . . and believe me, there are a lot of people who are rapping us for doing some of the things that you are doing, rapping us in the sense that we are cheapening the teaching process. A lot of the esoteric people in the field are saying, 'aren't you cheapening the teacher's role?' I happen to disagree with this 100%."

TA -- "There is one question we have touched lightly on. We are on an hourly wage. There is one great injustice that I wish you people would do something about and that's social security. Any job you get any place, scrubbing floors, just anything, you are covered by social security. Now we are supposed to be paraprofessionals, and we're not covered, and I don't think it's fair."

Another interesting finding of the project, and one related to the above premise, is the potential of teacher aide employment as a recruitment device for special education. Depending upon their background, experience, and competence, some aides sensing their potential may, with assistance, enroll in a progressively more demanding teacher preparation sequence leading to full teacher certification. Further, if the future development of paraprofessionals should lead to greater differentiation of role and function based upon level of competence (e.g., school aide, teacher aide, or school-community aide), the utilization of former teachers and those with only partial training in education may be facilitated by employment as a "teaching" or "instructional" aide. Conceivably some individuals with prior or partial backgrounds in education (or related disciplines) may prefer not to be responsible for the total scope and sequence of a program and may, instead, elect to function in a helping capacity to the professional teacher.

At the close of the first year of the project two of the aides (2 of the 12 or 16%) returned to school with the expressed intent of attaining credits for certification as teachers of the retarded. These aides were typically those who had already completed several years of college work. After the second year, at least 4 of the 12 aides (33%) in the project were enrolled in evening courses leading towards certification as a special education teacher. Thus, six project aides over a two-year period enrolled in a teacher-training sequence following exposure to special education as a paraprofessional.

PREMISE II: UTILIZATION OF PARAPROFESSIONALS IN SPECIAL EDUCATION PROVIDES EXCELLENT RECRUITMENT OPPORTUNITIES FOR TEACHER EDUCATION. FURTHER, IT OPENS UP AN UNTAPPED MANPOWER RESERVOIR OF PERSONNEL CAPABLE OF PERFORMING MORE LIMITED FUNCTIONS THAN A FULLY QUALIFIED PROFESSIONAL.

Positives

T -- "Special competencies of the aides have provided enrichment opportunities for the pupils."

T -- "If you are fortunate enough to have as your aide a teacher who quit because she had little children and now is coming back as an assistant, she can do a great deal of different type of things than the aide who has always been a housewife . . ."

A -- ". . . I would say that I think we should use all theknowledge and skills that we have at this particular time in hiring aides and in indicating to the teachers what particular skills the aide seems to have . . ."

A -- "Some aides, sensing their competence, may, with assistance, enroll in progressively more demanding programs leading to full teacher certification."

A -- "Teachers can and should assume a role in increasing the competencies of aides."

A -- "Studies show a high percentage of aides become interested in teaching and pursue course work along with their on-the-job training."

A -- "Proper use of aides is seen as a way of enhancing the professionalism of certified and qualified teachers. We are also finding that most schools are liberalizing their use of aides, permitting either officially, or more often unofficially, the use of aides in assisting with the instructional tasks and instructional-related tasks."

Negatives

T -- "One situation was temporarily somewhat volatile when a white teacher requested her black teacher aide to take some instructional materials out to her car. Other regular black teachers in the building observing this use of an aide expressed some resentment. The black on-site project coordinator was helpful in quickly 'cooling' the situation by indicating the limited background and training of the aide in question."

As one perceived by pupils as an adult authority figure in the school setting, the aide must be competent in reading, writing, speaking and computing. Beyond these fundamentals, the qualifications required of the aide may be determined by the tasks he is expected to perform. If special training is required, appropriate training programs should be developed by schools in which the aides will serve. Program leaders within the school system should be augmented by outside instructors from universities and elsewhere (Slagle, 1969).

Originally, the inservice thrust in terms of the paraprofessional working in the school situation was aimed at the teacher aides employed for the project. In fact, the political realities required the directors of the project to concentrate the training of aides to work with teachers. The agenda for the inservice training of the teacher aides, coordinated by the on-site project coordinator follows on the next page. Forty-four clock hours were involved in this orientation for the teacher aides. On the other side of the coin, training teachers to work with paraprofessionals received restricted emphasis in the early stages of the project.

The skills and personality of the teacher are critical variables to be considered in designating the initial role and function of the teacher aide. The ability of the teacher to implement an instructional program with the appropriate scope and sequence based upon pupil needs, and utilizing the proper instructional techniques and materials, are critical variables. The teacher's ability to translate her strategies to another individual, and his or her willingness or capacity to share their program with another, will greatly affect the manner in which an aide will be utilized. This is true even if a particular role and function has previously been designated for the aide. A teacher assigned an aide will need to develop, if they do not already have them, both administrative and supervisory skills if they are to properly and effectively utilize an aide ([illegible], 1968).

It soon became apparent that for a maximum payoff attention should be given to training teachers to work with aides as well as aides working with teachers. Thus, during the summer following the first semester of the project a special study institute was inaugurated focusing on better utilization of paraprofessionals by teachers in the project. Throughout the conference a continuing and recurring theme was the need for teachers who are going to have aides, and who have certain expectancies as to the aide's role, to be partners with local leadership personnel in orienting new aides to their functions.

A parallel concept repeatedly stressed at the conference was the thrust for a teacher of the 70's to function as an organizer, an expediter, or as a decision maker. In order to use a teacher aide effectively, it was thought necessary that the teacher first have a very clear vision of her own role as a teacher. She needs to understand the hierarchy of her own instructional tasks, and then decide which ones fully utilized her skills and training and which ones best involve her simply as an organizer, an expediter, or as a director.

It was further suggested that, in addition to the series of orientation training sessions for teachers and aides, it would seem well to plan at least one all-day session before the opening of school, right in the assigned school and with the assigned teacher and the principal. Most of the session would probably focus on school procedures to acquaint the aide with the school physical plant, schedule, entire school program, community, and to plan with the teacher. In this way the pre-service training of principals, teachers, and aides would be facilitated.

A further concern and felt need expressed by participants at the conference was the imperative need for teacher training institutions to incorporate understandings, skills and techniques in the use of paraprofessionals in the curriculum of special educators. Much of the original anxiety of teachers

INSERVICE TRAINING OF TEACHER AIDES

December 1, 1966	- 9 A.M. to 12 Noon The Role & Function of An Aide in the Special Class Setting	- Mrs. Sarah Scott Coordinator
December 6, 1966	- 9 A.M. to 12 Noon Interpersonal Relationships	- Mrs. M. Coulson School Social Worker
December 8, 1966	- 9 A.M. to 12 Noon Techniques & Methodology Utilized in Working with the EMR Child	- Mrs. Sarah Scott Coordinator
December 14, 1966	- 9 A.M. to 12 Noon Ethics Involved in Terms of Professional Information	- Mrs. Pauline Coggs Supervisor - Lay Workers
December 15, 1966	- Full Day - Tour of Goodwill Industries and The Ranch	- Mr. Bierman Mrs. Scott
January 4, 1967	- 9 A.M. to 12 Noon Use of Audio Visual Equipment Use of Ditto Machines	- Mr. Michael Swanson - Mrs. Lillian Miller
January 11, 1967	- 9 A.M. to 12 Noon Filling in Forms, Recreation for EMR Child, Blue Cross, S.S., etc., Observe Trainable Class	- Mrs. Sarah Scott Pleasant View School
January 18, 1967	- Lecture: Psychologicals and Slides on the EMR Child	- Dr. John Jackson, Supervisor Psychological Services
January 20, 1967	- Full Day Tour - Southern Colony	- Mrs. Sarah Scott Coordinator
January 27, 1967	- Organization and Placements	- Mrs. Sarah Scott Coordinator
February 13, 1967	- 4 P.M. to 5:30 P.M.	General Project Discussions
March 27, 1967	- 4 P.M. to 5:30 P.M.	General Project Discussions
May 22, 1967	- 4 P.M. to 5:30 P.M.	Subjective Impressions in Regard to Project, Also Turn in Evaluations

who were approached to participate in the project appeared to be related to lack of information and training with respect to their role as an expediter and as a decision maker in utilizing human and non-human resources. On the other hand, those teachers who had participated in the project were in agreement that investment in aide training, and ingenuity in utilizing the teacher aide, pays dividends in classroom interactions.

PREMISE III: APPROPRIATE AIDE TRAINING SEQUENCES SHOULD BE DEVELOPED BY THE SCHOOLS IN WHICH THEY WILL SERVE AND BY THE TEACHERS WITH WHOM THEY WILL WORK AUGMENTED BY QUALIFIED OUTSIDE CONSULTANTS.

Positives

T -- "I think that the actual value is really in proportion to the teacher's ingenuity in utilizing the aide and in her willingness to invest time in a little development of the aide's skills."

A -- "The paraprofessional should become familiar with the public schools as institutions, the professional responsibilities and ethics of teaching - as well as the major goals, special programs, procedures and policies of the school."

A -- "Teachers can and should assume a role in increasing the competencies of aides."

A -- "It rests upon the administration at that point to get teachers who apparently have the capacity to take on this rather critical responsibility. A lot of the training before and during will be inadequate or immaterial if the teachers themselves are not able to carry on the training aspect with the aide . . ."

A -- ". . . the qualifications required of the aide may be determined by the tasks he is expected to perform. If special training is required, appropriate training programs should be developed by the schools in which the aide will serve. Program leaders within the school system should be augmented by able outside instructors from universities and elsewhere."

A -- "Programs should include basic knowledge of school operations and procedures so as to enable the aide to become an integral part of the school staff. Also, certain skills for working with children may be identified as basic to all aides regardless of grade level or subject assignment."

Negatives

T -- "The question that keeps coming to my mind is, 'shouldn't the teachers who are going to have the aides then be much more a partner in the whole business so that the actual training of the aides is a combined type of operation?'"

T — ". . . in some way it should be made much more clear to the principal in charge as to what the aide's role is in the classroom . . ."

TA — ". . . but I would have felt better taking that class if I had been shown once in training about record keeping, about all those different types of records you have to keep, attendance and that sort of thing."

PREMISE IV: APPROPRIATE PRE-SCHOOL WORKSHOPS, INSERVICE, AND ON-THE-JOB TRAINING PROGRAMS SHOULD BE DEVELOPED FOR BUILDING PRINCIPALS AND TEACHERS WHO WILL BE UTILIZING PARAPROFESSIONALS. TEACHERS NEED TO DEVELOP NECESSARY ADMINISTRATIVE/SUPERVISORY SKILLS IN WORKING WITH AIDES AND BE ABLE TO CLEARLY CLARIFY THEIR INSTRUCTIONAL GOALS AND EXPECTANCIES FOR THE AIDE.

Positives

T — "I think you must remember that if the teacher can't relate to the aide what the goals and objectives are for the day, how are the children going to understand what they are supposed to be learning and what their goals are."

T — ". . . I like to think of a teacher aide as an investment, not a liability, but as an investment."

T — "In the area of planning we have to determine, first of all, what we're going to teach, to whom it will be taught, and which materials are going to be most useful. Once these technical level decisions have been made by the trained teacher, we can use the aide to extend the teacher's effectiveness in preparing the needed materials."

T — "One's whole day is going to be much more effective as far as teaching and it will require a higher level of achievement in teaching because one has been forced to clarify his goals and objectives to someone else."

T — "You do better planning, not just on a day-to-day basis, but on a long-range planning basis because this aide is going to be watching you all year as a teacher . . ."

T — "When you take on the responsibility of working with another adult in your classroom, you just have to be a more effective teacher."

A — "A pre-service training program which would be a joint training program for teachers, principals as well as aides is recommended. Previous studies indicate a vital factor in the success of employing and utilizing aides effectively is the ability of teachers and administrators to accept such a program positively."

A — "It tends to put 'starch' in their teaching when being observed by some other adult in the classroom"

A -- ". . . the teachers themselves then must get to be pretty sensitive to the learning problems of children because they, in turn, must communicate back to the aide so that the aide is sensitive to the kinds of learning problems that the child has."

A -- "It is recommended that inservice programs be established to acquaint principals and teachers with the types of assistance that can be obtained from the paraprofessional."

TA -- ". . . and it doesn't take long for the teacher and the aide to have more or less a certain set of signals, too, without even stopping . . . you know immediately what your teacher wants you to do . . . just by a lift of her eyebrow . . ."

Negatives

T -- "We need more opportunities to verbalize our concerns regarding the role and function of teacher aides and they with us."

T -- "I would say that a salary differential or allotment for the teacher who takes on an aide would give new insight into her role as one of training. If they take on this type of responsibility they should be compensated because they can be held more accountable for the type of training they performed."

T -- "I felt we should have had much more opportunity to relate to the people in charge of the program in terms of what we were finding as we worked along in these relationships and in conjunction with what they expected."

T -- "It should be made much more clear to the principal in charge as to what the aide's role is in the classroom and to the teacher and to the children . . ."

A -- ". . . and the teachers evidently did not specify even direction, but just left the issue hanging."

A -- "It was apparent that many of the teachers were unable or unwilling to answer the aides' questions about why they did things in a certain way and to explicate the goals that they were trying to achieve in doing certain activities in the classroom."

TA -- "I did feel it would have been helpful to have the project teachers and the aides get together occasionally and communicate rather than always having the teachers and aides meeting separately."

PREMISE V: TEACHER TRAINING INSTITUTIONS SHOULD BE RESPONSIBLE FOR INCORPORATING CONCEPTS AND SKILLS OF TEACHER-AIDE RELATIONSHIPS IN THE SPECIAL EDUCATION TRAINING SEQUENCES.

Positives

A -- "Continuing and timely training of aides will enhance their value to the school. On-the-job training should be planned and provided

A -- "I think the conclusion is rather obvious that we in the school business are going to have to find ways of making teachers go farther, some way of stretching them in the leadership roles, like the concept of the master teacher . . . and then augment them with many things: with hardware, with software, with teacher aides."

A -- "Teaching teachers to use aides effectively is the key to expanded teacher-aide programs. One need that so far is almost totally unmet is the inclusion of training in the use of aides in teacher education courses . . . since aides are a permanent and growing part of American education, it is essential that training for their use be provided . . ."

A -- "It is recommended that inservice programs be established to acquaint teachers with the types of assistance that can be obtained from the paraprofessional."

Slagle (1969) in a paper described the official position of the Wisconsin Department of Public Instruction as it relates to the use of teacher aides in the classroom. In this paper, he indicated that it is appropriate that the professional teacher be afforded useful resources, both human and non-human, to carry out his responsibility to meet pupils' learning needs. Boards of education may enable the use of teacher aides. Their actual employment is properly a professional determination involving teaching staff and administration. When boards of education have authorized the use of teacher aides, and certified faculty have elected to use them, administration should see that the lines of responsibility between aides and individual teachers or groups of teachers are clear.

The aide, then, is limited to assisting the certified teacher. He may not be assigned professional teacher responsibilities such as diagnosing educational needs, prescribing teaching-learning procedures, and evaluating the effects of teaching. He is regarded as an extension of the teacher, as an assistant who performs only at the bidding of the teacher. The aide does not reduce, remove, or transfer the certified teacher's authority or responsibility.

In order to facilitate the proper utilization of teacher aides in concert with the above-stated educational agency policies, Blessing (1967) provided the teachers and teacher aides in the project with an illustrative listing of teacher aide functions and types of duties. These projected functions follow:

Classroom

(1) Prepare classroom for use each morning; proper ventilation, lighting; put everything in proper order for the day - emphasize neatness.

(2) Keep bookshelves, reading tables, and chairs in reading circle neat and orderly; well-arranged books and workbooks.

(3) Assist children in keeping desks and storage places neat.

(4) Prepare for use and return to proper storage places needed teaching aids, such as: film strip projectors, screen, film strips, record player, etc.

(5) Keep teaching aids in neat storage and prepare for use when needed by teacher; return to orderly place when no longer needed - such as number games, word cards, etc.

(6) Assist in assembling, putting up, and dismantling bulletin board displays; keeping them neat, refreshed, and current.

(7) Maintain a neat classroom at all times.

(8) Keep chalkboards clean and ready for use.

(9) Assist in caring for plants, aquariums, etc.

Playground

(1) Supervise children passing from classroom to playground.

(2) Prepare and assemble materials for games, etc.

(3) Assist in physical education activities.

(4) Assist in supervision of playground with other teachers on duty; may relieve one teacher at a time.

Teacher

(1) Assist in use of filmstrip projector, filmstrips, record player, and other audio-visual aids.

(2) Consult with librarian and make it possible to utilize more fully the materials from the library in enriching a unit being taught.

(3) Assist teacher by helping small groups of children understand and follow instructions on seatwork while she works with another group uninterrupted.

(4) Assist in art classes in order that more media may be used; prepare materials in advance, such as mixing paint, cutting paper, sorting, and gathering materials.

(5) Assist in group and individual science experiments.

(6) Read stories for children during story time, rainy days, or when a child needs individual attention.

(7) Assist teacher in making files, keeping them in order, mounting pictures, etc.

(8) Maintain order if the teacher is called from the room for a few minutes.

(9) Assist the teacher in accompanying class on field trips.

(10) Make charts and word cards as directed by the teacher.

(11) Assist the teacher in individualized programs of study.

(12) Provide individual help and direction for programmed materials.

(13) Provide direction in concrete aids for mathematics.

(14) Help teachers put greater "goal specificity" into their instruction.

(15) Function as an "interpreter" for the school and the teacher clarifying objectives, rules and regulations. Very often the aides come from the same neighborhoods and backgrounds as the children and speak the same language as the parents.

Pupils

(1) Assist children in removing and hanging wraps; in reclaiming the correct ones when needed or at the end of the day; provide for neat storage of the same during the day.

(2) Assist in supervision of pupils passing to and from special classes such as music, physical education, library.

(3) Supervise bathroom breaks so that the teacher may continue with small-group work uninterrupted.

(4) Accompany individual children to the school clinic, dental clinic, etc., when necessary.

(5) Distribute and collect student papers.

(6) Assist with bus duty (at least one classroom teacher on duty).

(7) Observe children for indications of illness, emotional problems, physical defects, and extreme fatigue.

(8) Work with children during leisure and recreational reading activities.

(9) Work with children when special practice is needed after skills have been taught by the classroom teacher.

(10) Be aware of children's intellectual needs through conversation with them, listening to their stories and experiences. Children need good listeners.

(11) Help children utilize their independent work and play time.

(12) Aid in the development of oral speech and language by motivating, rephrasing and helping children use verbal mediators with which to think and speak.

Clerical

(1) Take roll call; make daily attendance report to principal's office.

(2) Collect and keep records on any special money, such as pictures, Junior Red Cross, etc.

(3) Assist in duplicating teacher-prepared materials for class use.

(4) Assist in such secretarial work as making pupil lists for eye tests, dental surveys, etc.

(5) Type tests and other materials already prepared by the teacher.

(6) Keep records of textbooks.

(7) Assist in keeping health records, weighing, and measuring children.

(8) Assist in checking out library books and keeping records.

Lunchroom

(1) Take lunch and special milk count; make daily report to cafeteria.

(2) Assist in supervision during lunchtime, helping children, encouraging them to eat well, but with another teacher also on duty.

These suggested roles and functions of a teacher aide were presented to the participating staff in the initial inservice sessions by the on-site project coordinator. Analysis of the on-going activities of the teacher aide, and subsequent discussion with participants on the project, indicated that the teacher aides had participated in a wide and diversified range of activities.

Our approach to the role prescription of the aide should be contrasted with the approach suggested by the local R and D Center for regular education. That is, the delineation of the role in our project was allowed to vary according to a host of interacting variables. It became our task then to specify who fit best with whom under what circumstances to achieve specified ends. This evidence then provides us with an empirical springboard from which to disseminate our findings and to provide the basis for a fruitful pairing of the professional and paraprofessional in the education of handicapped children.

The R and D Center, in contrast, has suggested an a priori subclassification of the aides' role into assistant teacher, instructional aide, supervisional aide, and clerical aide. The relative infrequency of special classes in most school systems precludes our being able to afford such restrictive role differentiations at this time. Also, the large number of people that were interested in becoming aides in the special classes in our project suggests that we might be able to afford to be selective in our choice of a more renaissance type of aide.

During the interim conference, and in the subsequent evaluation of the aide function, it became apparent that initially aides were utilized primarily in general housekeeping and clerical tasks in the classroom. However, as the teachers and aides became more familiar with one another's expectancies and competencies a trend was noted for the aides to become more involved in "assisting with the instructional tasks." In fact, some aides expressed some dissatisfaction if they were only permitted housekeeping or clerical roles and were not allowed direct involvement with children. Aides appeared to enjoy direct interaction with children rather than a preponderance of housekeeping tasks.

Since the range of aide duties varies from the mundane to actual instructional activity under the supervision of the teacher, it appears important to strike a balance between the two extremes. One of the consultants to our interim conference, Sister Sheila (1968), indicated this need in a similar vein.

> "The work that the aide gets to do should not just be a busy work sort of thing; the glorified secretary, or the glorified maid, or the glorified custodian. While it is true that the teacher would like to have many of these tasks relegated out of her hand, we think we have to show a certain balance in what we give the aide to do. The assignments should follow a certain pattern or routine. If they are definite and allow the aide to use a little initiative and ingenuity, she is going to be a lot more interested. The work of the aide, as much as the work of the children, must be planned for."

For some very meaningful and illustrative examples of teachers making technical level decisions and then permitting aides to function on an instructional basis, the reader is referred to Sister Sheila's paper, which is in Appendix A-3.

PREMISE VI: A NEED FOR BALANCE BETWEEN GENERAL HOUSEKEEPING, CUSTODIAL, CLERICAL AND SUPPORTIVE INSTRUCTIONAL ASSIGNMENTS IS IMPERATIVE FOR AIDE SATISFACTION AND SELF-ESTEEM. AIDES ENJOY DIRECT INTERACTION WITH CHILDREN.

Positives

T -- "The role and function of the aide changes the longer they stay in the assigned room."

T -- "Aides are more interested if the assigned tasks involve direct interaction with the children rather than a heavy preponderance of housekeeping jobs . . . Some of them [routine chores] the aide can take care of and is willing to take care of, but if that's all they ever get to do they might as well stay at home . . . I think we have to show a certain balance in what we give the aide to do."

T -- "My aide was able to help me in many situations of emergency with our children. In cases of illness or accident she was able to contact the parents, the school nurse, and the principal. Also, our social worker and our school psychologist came to our room many times and they were happy that we had an aide, because they were able to get her view of a particular child's behavior and achievement. I believe she was once asked to write a behavioral summary on a child . . . and this was quite a compliment to her because they felt she was capable of contributing to this particular situation."

A -- "Aides should be involved in planning and evaluating so that they may better understand their roles and ways to make maximum contributions."

A -- "Maximum use should be made of particular talents and interests so that aides may obtain the personal satisfaction which comes from maximizing their contributions."

TA -- "Because a master plan was made I knew just how far we went that day and I would then know where to begin the next day in case a substitute would come in. I could explain to her, you know, the situation of the day before and where the lesson plans let off . . ."

Negatives

A -- "Several of the project aides, after a year or so of experience, felt that they could do the job as well as the teacher they were working with. This raises an interesting question which could be directed towards teachers. That is, what is unique about their formal training which they feel helped them do a better job with the children as opposed to simply getting into the classes and by a 100% apprenticeship role becoming teachers?"

TA -- "I felt I was placed in an untrue instructional situation with a reading group by the teacher during certain filming in the project."

TA -- ". . . and I had to walk around in these tennis shoes during the filming to keep the noise down. . . to have a nice quite classroom. I think these kids should act normal and act as they do in a normal way. . . . I think the filming was a waste of money."

Another area of interest emerging from the discussions with the teachers and aides is the role of the aide in relation to the parents of the children. It will be recalled that in the previous listing of possible roles and functions for the aide that one function of the aide could be to serve as a community-agent or as an "interpreter" for the school and the teacher. It has been further suggested that since many of the aides come from the same cultural environment as the students, they can occasionally perform a social worker function without the stigma that social workers often bear. In other words, they work from the inside, not the outside. They talk education to their neighbors in their neighbors' language.

In this study some of the aides were rather actively involved with the parents except in the more formal parent-teacher conferences at which time they were excluded. Apparently, in their role as an interpreter for the school and the teacher, some aides assisted in clarifying the objectives, the policies, and the regulations of the school and instructor.

A number of the teachers reported that the school system's efforts to assist the children through the employment of teacher aides bettered the relations with a number of the parents who came in to work on special class projects and programs. Similarly, a number of the teacher aides indicated that parents had reported feeling more comfortable coming to school for a conference when an aide was present. School districts contemplating the employment and utilization of teacher aides should bear in mind the potential of the aide in serving the role of a school-community liaison person providing the essential communication contact between the school and the pupils' homes.

PREMISE VII: SOME TEACHER AIDES CAN FULFILL A VITAL ROLE AS A SCHOOL-COMMUNITY AIDE, PROVIDING LIAISON, AND COMMUNICATION BETWEEN THE SCHOOL AND PUPILS' HOMES.

Positives

T -- "I just want to say that in our group this past year we had two parents . . . who were outstanding members of the PTA . . . and they came up two or three times a month to work on special projects . . . I think this idea of the parents being more active was partly due to the teacher-aide project."

T -- "I think parents felt more comfortable when they came to a PTA meeting . . . more comfortable coming to school and they knew that just because they came they weren't going to hear their child criticized for not behaving in school . . ."

A -- "Among the various possible categories of paraprofessionals is the school-community aide who may provide the liaison between the school and pupils' homes."

A -- "The business of schools is teaching and learning and for a long time schools said to parents, 'Give us your children and stay home and don't bother us.' Parents pretty much accepted this and pretty much did just that. But times change and along with the times, the teacher's image of a parent and vice versa. There is now and needs to be much greater communication between the school and home. The aide is perhaps best suited to conduct the communication, bridging the gap between the parent and the teacher and freeing the teacher for more communication herself."

A -- "Public relations (home-school relations) may be improved by the addition of lay persons - especially those from the same cultural background."

A -- ". . . showing an effort as a school system to offer these kids more [aides] brought the parents to feel more comfortable."

TA -- "I think we got a very good response from the parents. Just something happened between us. I don't know why. It was something to do with the teacher setting the pace on the first day of school,

> telling the kids that I was a helper for them . . . So we just asked the parents to drop in . . . I got the feeling that the parents really appreciated me being there because if they would see me on the street . . . they just accepted me."

Negatives

TA -- "I don't think my situation was as successful as some of the other teacher aides. I think that I had a very good relationship with the children, but I never met the parents because I wasn't supposed to be there at the parent-teacher conferences."

In the taped film recording, teachers were requested to indicate a number of their satisfactions with respect to involvement with teacher aides in the project. It was interesting to note that all of the teachers in the project indicated their willingness to include a teacher aide in their classroom during the coming school year even though the project would no longer be in session. Initially the project planners had intended to employ aides trained in a private service agency in the local community. However, the on-site project director noted that the majority of these trained aides came from the upper socio-economic strata of the community and were typical of the volunteer-type individuals belonging to women's service organizations. She was quick to note that the majority of the teachers in the project came from the lower middle and upper middle classes and expressed her belief that these teachers would not feel too comfortable working with aides from the higher socio-economic levels. Therefore, a program of open recruitment for teacher aides was incorporated into the project.

On the question of teacher use of aides on a full- or part-time basis, the general reaction of the teachers was that a half-time aide would be most beneficial to them although a number of the teachers would have preferred to continue with an aide on a full-time basis. Several teachers expressed the need for some time by themselves with the children away from the observation and responsibility of an aide. As a matter of fact, after cessation of the project, the administrative office of the Milwaukee Public Schools assigned each of the aides to two teachers. The general reaction of the aides has been that they would function best by working with one teacher although the majority indicated their willingness to work with no more than two teachers.

Observations of successful teacher-aide teams leads to the conclusion that there must be some kind of chemistry taking place between teacher, aide, and children for the teacher-aide team to work. If the team is not functioning adequately, the process or the system has to permit readjustment of the situation to the satisfaction of all parties concerned. At least one teacher expressed her belief that positive interpersonal relationships between the teacher and aide were so critical that if the relationship was not working out satisfactorily, this situation needed to be rectified by the supervisor. In these circumstances, it was felt that the aide should be permitted to request a classroom transfer without penalty.

This being a research project, teacher aides were assigned randomly by the central office personnel. On the other hand, a number of the teachers concerned with interpersonal relationships expressed their belief that school districts should permit teachers a greater role in the interview process and selection of teacher aides for their specific classroom. No teacher should be assigned an aide against her will and a teacher should have the final say in the selection of the paraprofessional to be utilized in her particular learning environment.

PREMISE VIII: INTERPERSONAL RELATIONSHIPS ARE ENHANCED AND TEACHERS FEEL MORE COMFORTABLE DIRECTING THE WORK OF AIDES WHO COME FROM SIMILAR OR LOWER SOCIO/ECONOMIC LEVELS THAN THE TEACHER.

> This impression was gained from discussion with the on-site project coordinator. Over half of the classes participating in this research study were located in Milwaukee's core area.

PREMISE IX: NO TEACHER SHOULD BE ASSIGNED AN AIDE AGAINST HER WILL AND TEACHERS SHOULD BE PERMITTED A MORE ACTIVE ROLE IN THE SELECTION PROCESS. NOT ALL TEACHERS PREFER FULL-TIME RESPONSIBILITY FOR THE SUPERVISION OF AIDES AND SOME PREFER TO SHARE THE AIDE WITH ANOTHER TEACHER.

Positives

T -- "Now in all the conferences that we (teachers in the project) had I think we felt much more comfortable meeting separately . . ."

A -- "No teacher should be assigned an aide against his will . . . Selection of aides should accord with staff-felt needs . . . Teachers or other staff with whom aides will be working should participate in the decision to use aides and in the aides' selection procedure. Selection criteria should be agreed upon among personnel officials, administrators, and teachers."

A -- "The decision to use aides must rest primarily with teachers . . . He is regarded as an extension of the teacher, as an assistant who performs only the bidding of the teacher. The aide does not reduce, remove, or transfer the certified teacher's authority or responsibility."

Negatives

T -- ". . . I just happened to remember a comment I heard made by someone and this person felt that they would have loved to have had their class without the aide just one day or so at the start of the school year. She could have probably established a rapport with the children before the aide came because it was difficult after the aide came to establish this kind of working relationship, if you know what I mean."

A -- "A number of the teachers felt that a half-time aide would be most beneficial to them. These same teachers expressed the occasional need for some time alone by themselves with the children away from the constant surveillance of the aide."

TA -- ". . . because I think the teachers felt freer to be very open about the criticisms as well as the good points and the aides likewise. But I do think . . . when they got accustomed to each other that all the meetings should then be with teachers and the aides together. At first both are feeling their way in the new situation."

PREMISE X: TEACHER AIDES PREFER WORKING WITH ONE TEACHER AND THE MAJORITY FEEL THEY COULD WORK FOR NO MORE THAN TWO.

Positives

A -- "The general reaction of the aides has been that they could function best by working with one teacher."

TA -- "I can say personally that I could not work for more than two teachers. I don't think I could, maybe I'm too one-track minded, but I don't see how you can really get to know and work well with more than two people plus all the children."

PREMISE XI: INTERPERSONAL RELATIONSHIPS ARE SO IMPORTANT IN SUCCESSFUL TEACHER-AIDE SITUATIONS THAT THE SYSTEM MUST PERMIT READJUSTMENT AND REASSIGNMENT OF AIDES WHERE NECESSARY.

Positives

T -- "Inasmuch as interpersonal relationships play a major role in the teacher-aide relationships, I think they should be made aware at the beginning of each situation that should these particular situations not work out they should feel free to report this to whomever is in charge . . . so the situation can be rectified in some sort of way with no stigma attached to teacher or aide."

T -- "Teachers have the preference for putting in for a transfer if they wish within the year to go to a different school. I think this should be the prerogative of the aide if she is very unhappy in a situation. She should also make a request and it should be honored as far as you can."

T -- ". . . there still has to be some chemistry taking place between these two people in that room with those kids and after a period of time we should be able to say to somebody, 'Look, this is what happened to us . . . . this is how we feel . . . this is where we are now' . . ."

A -- ". . . I think any system should build in flexibility to make reassignments of responsibilities and of staff."

Negatives

A -- "It was apparent that a number of teachers defined the aide's role in such a way that their [teachers'] status was upheld in the eyes of the kids. This usually resulted in a negative presentation of the aide and the result appeared to be that the aide wasn't making a maximal contribution to the education of the children."

TA -- "When sitting in on these discussions, I find the biggest problem is my concern over what the teacher thinks of me, e.g., as just an aide, and how the faculty accepts me. I know everybody has some problems."

Reference has been made to the need for the school system, the principal, and the teacher to specify clearly the role and function of an aide. Contrucci (1968) has suggested that when an aide is designated to have primary responsibilities for secretarial, housekeeping, or instructional duties, or as a mother surrogate, or is given a combination of roles and functions, this sets some type of direction for the teacher and aide to develop their relationship. The physical and emotional environment in which the teacher, aide, and pupils interact, both from the building and classroom standpoint, can be important variables in the determination of the success or failure of a teacher-aide situation. A number of the teacher aides expressed some concern with failure on the part of teachers and administration to clearly delineate their limitations of authority. For example, in the matter of student discipline one aide expressed the need for a clear understanding of the limits of her authority in this area. She had noted through experience that she could be quite helpful when children in the special education class were "working over" an inexperienced substitute teacher. There had been a felt need for more training in control and discipline of pupils, but on the other hand there was some concern on the part of the aide because of a lack of a clear delineation of her authority role.

It was apparent that a number of teachers had defined the role of the aide in such a way that the teachers' status was upheld in the eyes of the students. This usually resulted in a negative presentation of the aide and the results appeared to be that in some instances the aide was not making a maximum contribution to the education of the children. Again there was the notion of uncertainty about role limitation. In a number of instances the aide did not know how far she could go or could not go in disciplining, instruction, etc., and the teachers evidently did not specify even direction, but just left the issue hanging.

A number of the teachers and teacher aides continually stressed the notion of the aide playing a "back-up" role in support of the teacher and her authority. This was particularly true in situations external to the classroom, on the playground, on field trips, and with substitute teachers. As one aide indicated, when working in a small group situation she had been given authority to send pupils to the bathroom and to supervise their rapid return to the group situation. This illustrative kind of situation clearly indicates how lack of agreement on the part of the teacher and aide could conceivably create little frictions in the teacher-aide interaction.

Another recurring problem expressed by several of the teachers appeared to be that of having to share their children with another adult authority figure. This seems to be an issue which could be explored with some benefit to education generally. It would appear that it would be desirable to be able to define the teacher as either a facilitator or expeditor of the academic instructional process or as a mother surrogate. This is not to say that these two roles are necessarily antagonistic, but it would seem that one or the other should have priority status with respect to teachers and with teacher aides.

PREMISE XII: THE SCHOOL SYSTEM, PRINCIPAL, AND TEACHER MUST CLEARLY DELINEATE AND SPECIFY THE EXPECTANCIES AND LIMITS OF THE AIDE'S AUTHORITY IN CONTROL AND MANAGEMENT OF PUPILS AS WELL AS IN THE INSTRUCTIONAL REALM.

Positives

T — ". . . what develops is you go into a program like reading . . . you would introduce the new vocabulary and a little bit of phonic work with the children . . . then I would hand the group over to Mrs. ____ who at this time would start working with the children. The more the aide showed effectiveness in working with these children, the more I would ask her to work with these kids. A great deal of involvement came as she showed more experience with the material and that she understood it better. And the more I would ask her to do . . ."

A — "After discussing it with the teacher, a decision is made as to who and how to handle the problem. Many times the aide can handle it herself."

A — "Many varied forces and factors influence both the manner in which a teacher aide is utilized and the success of an aide in a particular program . . . There are institutional goals and variables, teacher goals and variables, and aide variables. Of paramount importance are the particular goals and objectives which the educational institution has set as outcomes expected thru the utilization of aides . . ."

A — ". . . administration should see that lines of responsibility between aides and individual teachers or groups of teachers are clear."

A — ". . . the efficacy of the use of teacher aides is determined by observing the accomplishments of the school objectives . . . evaluation should be in terms of those objectives.

TA — ". . . this should be clarified, e.g., the aide should have the responsibility of letting children go to the bathroom . . . She should be able to send them off since she is working right there with them. I mean all these little things are very important. They really are, because they can cause trouble in the classroom."

Negatives

A -- "There's the matter of responsibility, and then there's the whole matter of authority and these are not the same . . . You may be given the responsibility for some activities and may not be given the authority for the activities . . . like discipline for example. You're not quite sure what authority or responsibility you have in the matter . . . it may be that these lines haven't been clarified for you."

A -- "Also related to this problem of role definition was the notion of uncertainty about role limitation. In a number of instances the aide did not know how far she could go or could not go in disciplining and in instruction . . ."

TA -- "I think if the aides had more training in disciplining and how far they should go in disciplining and who stands with them . . . this would be of great help to us because quite often you are caught and the teacher isn't around . . . you have to do something . . . and sometimes when a substitute is there you have to do something and not knowing how far you can go or how much to do, or not being trained, it sort of leaves you out on a limb."

TA -- "That is one of the most frustrating things. You have to work sort of spontaneously, anything they tell you to do. You do one job and you don't know what to expect next. I am one of those type of people that have to know what's coming off next . . . otherwise you finish your chore and the teacher may be in the middle of a reading group. You have to sit there and wonder what you're going to do and you feel that you're wasting your time . . . This is very frustrating to me."

PREMISE XIII: THE TEACHER'S ROLE AS A FACILITATOR AND EXPEDITER OF THE LEARNING PROCESS SHOULD BE CLEARLY DEFINED.

Positives

T -- "Having an aide provided me opportunities to get into more areas . . . there are some very positive aspects to this thing in terms of what you can do for the children which are not necessarily traditionally seen as in-classroom kinds of things . . ."

T -- ". . . she sees the role of the teacher in three segments: organizer, expediter, and director. The last of these three seems most important, especially when using an aide, but in each of the three the aide can spell the critical difference between pupil control and pupils out of control."

T -- "A study conducted at St. Coletta's found that the pupils were wasting nearly 40% of their time in the class. Here the teacher as director could instruct the aide to fill in that time with something useful to the student."

A -- "Let me talk a little bit about this business of being an expediter. In 10 to 20 years, special classes will be playing another role, a role they should be playing now, and that is acting as a center of referral . . . an agency responsible for this particular program which is concerned with the total programming needs for the child and his family. This is one of the major roles we should give to the classroom teacher, our master teacher."

A -- ". . . and the teacher must also be an individual who can make decisions because she will constantly be having to make decisions regarding the interaction of the aide with the pupils and the kinds of responsibilities the aide will assume."

A -- "Making decisions about pupil needs and ways to most effectively meet them is a primary professional responsibility of teachers. Careful analysis of professional responsibilities enables teachers to separate those requiring professional initiation and performance from those requiring only professional supervision . . . the essential point is that the teacher, as a professional, is expected to make decisions about pupil learning needs . . . appropriate resources may include teacher aides."

TA -- "We always followed through on that with her giving me her plan and I would find out exactly what she was going to do . . . and I'd go right ahead and do it."

Negatives

A -- "It was apparent that many of the teachers were unable or unwilling to answer the aide's questions about why they did things in a certain way to explicate the goals that they were trying to achieve in doing certain activities in the classroom."

PREMISE XIV: THE USE OF AN AIDE TO COPE WITH LARGER CLASSES UNDER CERTAIN CIRCUMSTANCES DOES NOT APPEAR TO BE AN UNREASONABLE CONCEPT.

Positives

T -- "In response to a question regarding continuance with an aide and an increased class size, one teacher stated: 'Well, I had positive results with my aide and the 22 children didn't make that much difference in discipline. The children who were unruly, etc., before the aide came, remained unruly after the aide came, you see, so I would say 'yes I would.'"

A -- "One of the purposes of utilizing an aide is to cope with large class size, shortage of teachers, and shortage of classrooms."

A -- "The Department of Public Instruction does not advocate the use of aides to displace certified personnel. However, it is recognized that the increasing assumption by aides of non-teaching functions formerly performed by certified teachers may well lead to reconsideration of acceptable ratios between certified teachers and pupils."

A -- "A number of teachers felt that an increased class size would pose no problem providing they were allowed to have the say based upon the existing complement of the class. The matter boils down to . . . the problem of the teacher's individual capacity to cope with the demand, plus the amount of supportive personnel and/or facilities available on the scene."

A -- "Size of special education classes appears to have resulted from a combination of teacher-administrator compromise, embellished with psychological justification and tradition."

A & T -- Note: When this project was initiated, state department policy required an enrollment increase of approximately 50% over the suggested ranges for classroom units under the supervision of a single instructor. Empirical experience with the use of aides in the field, and in this project, led the state educational agency to modify this requirement in 1967-68. Although some enrollment increase is still encouraged, situational variables at the local level call for individual negotiations with the field supervisor on the question of class size with an aide.

Further, in a post hoc conference with the on-site project director, state agency officials were confirmed in their belief of the necessity of the above-mentioned policy change. The on-site director indicated that no teachers in the project would accept an aide if a flat 50% enrollment increase were required, whereas the majority would accept some increase in class size depending upon the individual classroom circumstances and their involvement in the process.

Negatives

A -- "An additional purpose for introducing a teacher-aide program into special education has been noted - to increase the number of mentally retarded per teacher . . . It is interesting to note that this concept, which was evidenced only in the literature concerning special education, seems in substantial contradiction to the position advanced by the National Education Association . . ."

CHAPTER VI

CONCLUSIONS

The conclusions to this study can be grouped under three headings: (1) limitations of the study, (2) implications for administrative decision making, and (3) suggestions for future studies.

A. Limitations of the Study

The major limitation of the study is that the results can only be generalized to other EMR populations of pupils and to other areas and administration structures with a great deal of caution. The limited success in the attempted representative sampling makes generalization to Milwaukee as a whole hazardous, and to other urban and rural settings even more so. However, in the face of having to make decisions it is better to have input of a restricted nature rather than none at all. The review of the literature indicated that no quantitative data existed regarding the effect of teacher aides on EMR pupils' performance in the classroom. Studies examining the effect of class size were equally bereft of significant data.

Another limitation could be the length (three semesters) of the study. That is, would the impact of the aides have been more obvious had the study gone on for an additional period of time? Two factors suggest this could have happened. First, the analysis revealed a slightly greater (yet insignificant) increase in the slope of academic scores for classes with aides; secondly, it was evident that the teachers were "learning" how to use aides during the first and part of the second semester, thereby allowing the impact of the aide to be minimal during this period.

An additional consideration in a study of this type is the impact of the personal and probably ideological forces at work in the community at the time the study is conducted. Between the first and third semester (summer and fall, 1967) of the project racial strife and rioting was rampant in the area in Milwaukee in which the study was conducted. That this upheaval should not be reflected in the behavior of the pupils is unlikely. Indeed, one could hypothesize that much of the marked increase in the Disruptive Behavior scores of the BAIC between the first and second testing of the study could be attributed to the state of the community at this time.

B. Implications for Administrative Decision Making

The major impact of the aides alone (i.e., without regard to other conditions such as size of class, age level of children, etc.) appeared to be, first, a reduction in undesirable behavior in the classroom and, second, an enhancement of the spelling achievement of the pupils.

When questions of cost per unit of academic achievement are raised, the following recommendations can be made: either increased size without an aide, standard size without an aide, or increased size with an aide, (For all these conditions the average cost per pupil per 10 months of academic achievement = $2,343) are to be preferred to the standard size class with an aide (average cost per pupil per 10 month achievement = $4,033).

Class size alone had an effect on both arithmetic and classroom behavior. The standard (i.e., smaller) classes gained more in arithmetic and manifested less negative behavior than the increased classes.

In general, what effects the factors of aides and class size had appeared to be most pronounced in the primary classes suggesting that administrative strategies other than the use of aides and manipulating class size should be developed for the intermediate EMR pupils in the areas of academic, behavioral, and linguistic functioning.

Finally, it behooves school administrators to effect a fine balance between the primacy of either function or structure in the utilization of teacher aides. Too much structure, in terms of rigid role definition, teacher-aide, aide-parent and aide-pupil relationships, etc., will nullify the full contribution that the aides can make. On the other hand, the newness of the situation to aides dictates that some broad guidelines and training opportunities need to be provided.

C. Suggestions for Future Studies

Because of limitations in sampling in this study, replication is suggested to test the generality of the findings.

The Behavior Analysis Scale (BASE) developed in the study, requires examination in terms of its relevance as a monitoring device or merely as a pre-post research instrument. Preliminary evidence has suggested the BASE may have further value as a device for matching teachers and pupils. That is, certain teachers apparently cannot tolerate (as manifested by consistently high scores) certain of the five behaviors indexed by the instrument.

The language analysis procedure requires further study as a possible quick, easy, and non-reactive evaluative device for the many educational projects now being funded under federal auspices. In addition, developmental data should be obtained from other groups of children at different age and IQ levels.

Finally, many of the relationships which could be explored using the data were not. Considering the cost involved in collecting such data and its relevance to ongoing concerns in the education of the EMR, an attempt should be made to exhaust the additional information contained in the data.

REFERENCES

Anastasiow, N. *Project II: An Evaluation of the Kindergarten Teacher Assistant Project.* (Report of Title I Project Number 43-170-62-0124). Washington, D.C., U.S. Department of Health, Education and Welfare, Office of Education, 1966.

Blessing, K. Use of teacher aides in special education: A review and applications. *Exceptional Children*, 1967, *33*, 107-113.

Bock, R.D. Contributions of multivariate experimental designs to educational research. In R.B. Cattell (Ed.) *Handbook of Multivariate Experimental Psychology*, Chicago: Rand McNally & Co., 1966.

Bock, R.D. and Haggard, E.A. The Use of Multivariate Analysis of Variance in Behavioral Research. In Dean Whitla (Ed.) *Handbook of Measurement in Education, Psychology, and Sociology*. Boston: Addison-Wesley, 1968.

Bonn, M. Have traveled, will teach: Peace corpsmen come home with ideas. *American Education*, 1967, *3*, 1-3.

Contrucci, V., Utilizing the teacher aide, *The Amplifier*, March, 1966.

Cruickshank, W.M., and Haring, N. *A Demonstration: Assistants for Teachers of Exceptional Children*. Syracuse, N.Y.: Syracuse University Press, 1957.

Edelfelt, R.A. A teacher and his staff. *Virginia Journal of Education*, 1967, *60*, 11-13.

Editorial. Teachers (adult) helpers. *Scholastic Teacher*, 1967, *51*, 2 & 4.

Emmerling, F.C. and Chavis, K.Z. The teacher aide. *Educational Leadership*, 1966, *24*, 175-183.

Finn, J.D. A general univariate and multivariate analysis of variance, covariance, and regression model and program, *Behavioral Science*, 1969.

Goldstein, D.B. Teacher aides: The Indianapolis plan may lend itself to your school. *The Instructor*, 1965, *75*, 123-124.

Goodman, L., and Arnold, I. Training and utilization of nonprofessional personnel in services for the retarded. Mental Retardation, 1967, *5*, 11-14.

Gordon, L.V. *Survey of Interpersonal Values*, SRA, Inc., 1960, Chicago, Ill.

Greene, D.G. The value of an attendant in the classroom for the trainable mentally handicapped. *National Catholic Education Association Bulletin*, 1966, 63, 472-473.

Howell, J. Teacher Assistants: The Yale-Fairfield Study of Elementary Teaching. New Haven, Conn.: Yale University, 1959.

Jastak, J.F. and Jastak, S.R. WRAT Manual, Guidance Associates, Wilmington, 1965.

Kanner, L. History of the Care and Study of the Mentally Retarded. Springfield, Illinois: Charles C. Thomas, 1964.

Kirk, S. and Johnson, G.O. Educating the Retarded Child. New York: Houghton-Mifflin Co., 1951.

MacLennan, B.W. New careers as human service aides. Children, 1966, 13, 190-194.

McCarthy, J.J. and Kirk, S.A. Illinois Test of Psycholinguistic Abilities: Examiners Manual. University of Illinois, Institute for Research on Exceptional Children, Urbana, Ill., 1961.

National Education Association. Teacher aides in the public schools, NEA Research Bulletin, 1967, 45, 37-39.

National Education Association. How the profession feels about teacher aides; Teacher opinion poll. NEA Journal, 1967 56, 16-17.

Nelson, G. S 721, Teacher Aide Program Support Act of 1967. National Elementary Principal, 1967, 46, 40-41.

Lamp, A.C., and Crossen, F. Teenage teacher-aide project. National Elementary Principal, 1967, 46, 43-48.

Levenson, D. Mothers bring their skills. Parents, 1967, 42, 64-66.

Office of Compensatory Education. California Plan for the Education of Migrant Children. Sacramento: California State Department of Education, 1968.

Riessman, F., and Gartner, A. Paraprofessionals: The effect on children's learning. Urban Review, 1969, 4, 21-22.

Rankin, P.T. Evaluation of the Bank Reading Demonstration Project, 1966-1967. Detroit: Detroit Public Schools, Research and Development Department, 1968.

Research and Development Center for Learning and Re-Education. Report on Teacher Aide Program. Madison: University of Wisconsin, University Extension Bookstore, 1967.

Research Division - NEA. Class Size. Washington, D.C.: National Education Association, 1968.

Scheffé, H. A method for judging all contrasts in the analysis of variance, Biometrika, 1953, 40, 87-104.

Sheila, Sister M. Paraprofessionals in the school - I. The role of the teacher aide in pupil management. Bureau Memorandum, June, 1969.

Shields, E. Teacher assistant to the rescue. School and Community, 1966, 53, 31.

Shipp, M.D. Teacher aides: A survey. National Elementary Principal, 1967, 46, 30-33.

Siegle, A. The Teacher Aide: The Position of the Department of Public Instruction (Pamphlet), Madison: Wisconsin State Department of Public Instruction, 1969.

Stroever, R.C. AID program write-up social, sciences data and program library service, University of Wisconsin, Madison, 1970.

Teacher stretchers. American Education, 1967, 3, 27-28.

Thomas, E.L. Listening mothers. The Instructor, 1967, LXXVII, 20-21.

Thorndike, R. Pygmalion in the classroom: A review in American Educational Research Journal, 1968, 5(4), 708-711.

Valdes, D.F. Noon-duty assistant program. NEA Journal, 1964, 53, 63.

Webb, E.J., Campbell, D.T., Schwartz, R.D., and Sechrest, L. Unobtrusive Measures: Nonreactive Research in the Social Sciences, Rand McNally & Co., Chicago, 1966.

Williams, E.F. Tomorrow will be different, Virginia Journal of Education, 1966, 60, 16-19.

101 Tasks for the Teacher Aide. Croft Educational Services, 1968.

APPENDICES

APPENDIX A

Articles and Papers

1. A Look at Some Assumptions in Special Education
2. Teacher Expectancy and Student Achievement
3. The Role of the Teacher Aide in Pupil Management
4. The Utility of Mental Age as a Predictor of Academic Achievement
5. Noise Levels in the Classroom
6. Light Levels in the Classroom

JOHN J. COOK
ROBERT H. SCHWARZ

A Look at Some Assumptions in Special Education

Abstract: A posteriori analysis of data collected on EMR pupils allowed the scrutiny of a number of commonly held beliefs or assumptions about the educational process. These assumptions pertained to the regression phenomenon over summer, the beneficial effect of summer school, the negative effect of changing teachers, and the negative effect of high absenteeism as reflected in academic achievement. The social factor as it pertains to academic achievement was also examined. The findings suggested that all these assumptions are extremely suspect and that more closely controlled experimentation should be undertaken to examine these and other assumptions which have been guiding educational endeavors.

There are a number of assumptions made by educators which have been subject to little, if any, examination. Since these assumptions influence decisions about program development, they need to be examined. The present paper presents the results of the examination of four of these assumptions as they relate to the educable mentally retarded. The four assumptions are as follows:

1. There is a regression or loss of academic skills over the summer.
2. Summer school programs have a positive effect on academic learning.
3. Changing a child's teacher during the school year has an adverse effect on the learning of academic subjects.
4. The more a child is absent from school, the less he will learn of academic subjects.

A continuing search of research literature has failed to reveal any research pertaining to these assumptions as they relate to the EMR. This lack would seem to suggest one of three things: (a) the assumptions are self evident, (b) the assumptions are continually supported by the observations of the personnel involved, or (c) it is meaningless to undertake research studies in the four areas because of the number of other variables which influence learning in children.

In examining the four assumptions, data were used from a study having somewhat different concerns. The analyses of the data were post hoc, and subject to all of the restrictions in this type of analysis.

Data Collection

During the course of a project concerned with determining the differential effect of class size and use of aides on the achievement of EMR pupils, a number of dependent measures were obtained over a period of three semesters, academic achievement was evaluated. The dependent measures used were the reading, spelling, and arithmetic subtests of the revised Wide Range Achievement Test (WRAT).

The pupils were tested on three occasions: February, [illegible]; September-October, [illegible]; and May-June, [illegible]. Considerable pupil movement in and out of the project occurred during the three semesters. However, a total of [illegible] students remained throughout the project and were tested on all three occasions. The [illegible] of the various samples entering into each analysis are therefore presented. Procedures and results are stated separately for each of the four assumptions being tested. The statistical null hypothesis is implicit in testing the assumptions, with differences significant at the [illegible] level for a two-tailed test being acceptable.

Procedures and Results of Analysis

Assumption One: During the summer months EMR pupils lose in academic achievement.

Procedure: Of the [illegible] pupils tested in February, [illegible] were selected randomly for retesting in June. Routine project testing was done the following September. [illegible] of the [illegible] pupils were still in the project at this time and their data were used to test the assumption.

Results: Characteristics of the pupils as a function of age grade placement are shown in Table 1. The primary and intermediate classes did not differ significantly in terms of IQ scores, sex, or race. A repeated measures t test was done on the three WRAT subtests. Since the age factor could account for the differences observed, separate analyses were made for the primary (ages [illegible]), the intermediate (ages [illegible]) classes, and for the total group (see Table 2). Significant change was noted for only the reading scores. The group mean showed an increase of [illegible] raw score points over the summer with the primary group gaining [illegible] points and the intermediate gaining [illegible] points. This differential growth has been noted in other data obtained during the school year and suggested that the intermediate group's academic achievement was starting to accelerate rather rapidly.

An obvious explanation for the continued growth in reading during the summer months was that the children attended remedial classes during the summer. This assumption is explored in the next analysis.

Assumption Two: EMR pupils who attend summer academically oriented classes should demonstrate greater achievement over the summer months than those EMR pupils who do not attend summer school.

Procedure: Of the [illegible] pupils tested prior and subsequent to the summer months, [illegible] remained in the project throughout the three semesters. Eight had attended summer school and the remainder ([illegible]) had not. Analyses of the data consisted of a graphic display of the subtest scores over the four testing

TABLE 1

Characteristics of Pupils Tested Before and After Summer Months

	N	C.A. (mos.)	I.Q.	% Male	% Female	% White	% Non-White
Primary	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Intermediate	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE 2
Significance of Pre–Post Summer Differences of Three WRAT Raw Subtest Scores for Primary, Intermediate, and Total Groups

Reading

	N	Pre X̄	Post X̄	diff	t	p
Primary	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Intermediate	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Spelling

	N	Pre X̄	Post X̄	diff	t	p
Primary	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Intermediate	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Arithmetic

	N	Pre X̄	Post X̄	diff	t	p
Primary	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Intermediate	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Total	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE 3
Characteristics of Pupils Attending and Not Attending Summer School

	N	C.A. X̄ (Months)	IQ X̄	% Male	% Female	% White	% Nonwhite
Attended Summer School	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Did not Attend Summer School	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

assumption. Practical rather than statistical significance is the important consideration in testing this assumption.

Results. Salient characteristics of the pupils who attended and did not attend summer school are shown in Table 3. Differences between the two groups were minimal except for the sex composition of the two groups. Of the pupils who attended summer school, [illegible] percent were female. Of the group who did not attend summer school, only [illegible] percent were female. The predominantly [illegible] composition ([illegible] percent) of both groups reflected the composition of the project sample. The teachers who participated in the project were volunteers. Many of them taught in schools in the inner core of Milwaukee.

Figure 1 displays the mean scores on reading, spelling and arithmetic respectively for EMR pupils who did and did not attend summer school. A number of consistent findings stand out from the graphs. On all three subtests, the children who attended summer school were higher achievers. This could be attributed to the greater number of females in the summer school group or to the parental pressure to achieve in this group. Studies with regular grade pupils have consistently demonstrated that girls in this age range do on the average achieve better than boys. The higher achieve-more motivation of the summer school students and their parents can be assumed to be the nature of the summer school system, since summer school is voluntary and parents must provide transportation.

Finally, it should be noted that in the three subjects the differences between the two groups diminish markedly over the three semesters. This would suggest that the long term effect of attending summer school is negligible, or, possibly, negative.

Even more startling is the short term effect of attending summer school. In reading, the summer school group stayed at a score of [illegible], whereas the no summer school group went from a score of [illegible] to [illegible], a gain of [illegible] points. At this level on the WRAT in reading two points is equivalent to about a month of academic achievement. It would seem, therefore, that the pupils who did not attend summer school continued their academic growth in reading at the rate of 1 month in 3 during the summer. This compares with 3 months in 4 of the previous summer and [illegible] months in 8 in the succeeding two semesters.

The immediate effect of summer school on spelling was equally negative but in a somewhat different way (see Figure 1). The no summer school group had a level of [illegible] points on the average, whereas the summer school group went from a score of [illegible] to [illegible], a loss of [illegible] points. At this level on the WRAT, [illegible] points is equal to about one month's growth, the amount the summer school group lost.

In arithmetic, both groups maintained their levels over the summer (see Figure 1) and proceeded to grow academically at the rate of 3 months in the next 8.

Assumption Three. Pupils who have one or more teacher changes over a period of academic learning will achieve less than those who have had the same teacher.

Procedure. Of the [illegible] EMR pupils who were tested on the three occasions, [illegible] had had a change in teachers at least once. Only four had a change more often than once. Ideally, the best way to test

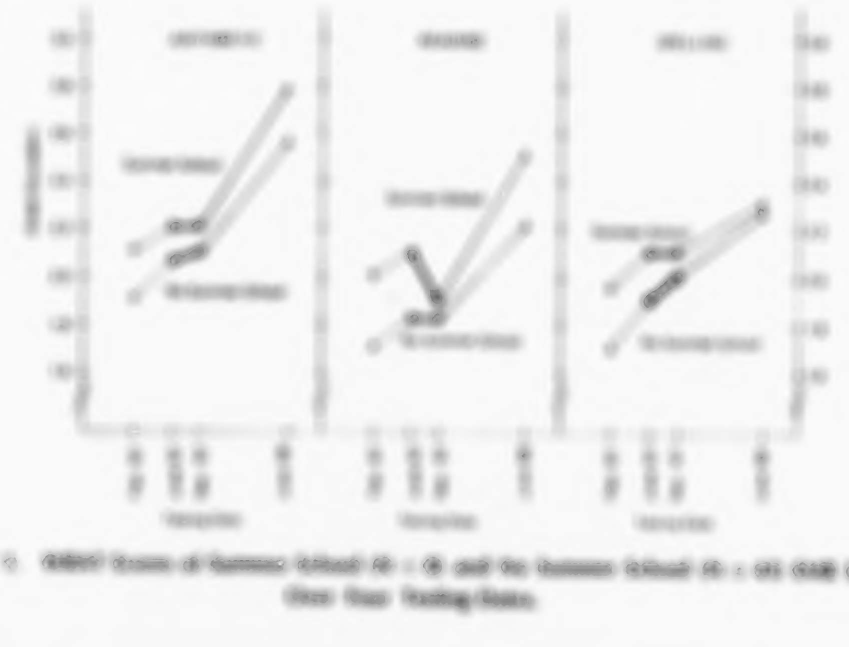


FIGURE 1. WRAT Scores of Summer School (N = [illegible]) and No Summer School (N = [illegible]) EMR Students Over Four Testing Dates

this assumption would be to have a large enough group of multiple change pupils, thereby maximizing the possibility of detecting differences due to the change factor. This was not possible in the present study.

A random sample of [illegible] pupils was chosen from the pool of [illegible] pupils who had not changed teachers. A repeated measures analysis of variance was done. The effect which was crucial to a test of the assumption was the treatment (change and no change) x time effect.

Results. Table 4 shows the characteristics of the change and no change groups. The only significant difference was the racial composition, with the non-white tending to characterize the change group more so than the no change group. An F ratio of less than one for the time x treatment interaction clearly indicated that the fact of having or not having the same teacher did not affect the pupil's achievement in reading, spelling, or arithmetic over the three semesters covered by the project.

Assumption Four. EMR pupils who have a high absence rate achieve less than pupils with a low absence rate.

Procedure. One hundred and forty-eight pupils had been given the WRAT on the three occasions of project testing. Absence rates had been obtained on all pupils for each semester. An average of the three absence rates was obtained for each pupil, and the 148 pupils were ranked in terms of frequency of absence. The 37 pupils in the upper and lower quartiles were designated the low and high absence rate groups. The median rate for the high group was [illegible] percent, while for the low group it was [illegible] percent.

Results. Table 5 shows the characteristics of the high and low absence rate groups. No significant differences were detected between the groups in terms of C.A., IQ scores, race, or sex. As expected, the time effect and testing occasions were significant on all three subtests. The low absence rate group was characterized by a significantly higher ($Fs > [illegible]$, $df = [illegible]$) average reading ([illegible] vs. [illegible]) and spelling score ([illegible] vs. [illegible]) than the high absence group. No significant differences were noted in arithmetic score. The lack of a significant interaction between testing occasions (time) and treatment (group) indicated that the rate

TABLE 4

Characteristics of EMR Pupils Who Had Changes in Teachers and Those Who Did Not

	N	$\bar{X}$ C.A.	$\bar{X}$ IQ	% Male	% Female	% White	% Non-White
Change	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
No Change	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

TABLE 5

Characteristics of Low Absentee Rate and High Absentee Rate EMR Pupils

	N	C.A. $\bar{X}$ (Mos.)	IQ $\bar{X}$	% White	% Non-White	% Male	% Female
Low Absence Rate	37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
High Absence Rate	37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

of academic growth over the three semesters was the same for both the high and the low absentee groups. Since we know that primary pupils' rate of achievement was higher than the achievement rate of intermediate pupils, the composition of the two groups in terms of this variable was informative. The high absentee group had 41 percent primary pupils and 59 percent intermediate pupils, while the low group had 51 percent primary pupils and 49 percent intermediate pupils. These percentages were not statistically significant ($t = 1.09$, $p > .05$).

Achievement as Related to Race. Over the past several years there has been a growing interest in differences in achievement as a function of race. This concern is explored in this section as it applies to our EMR sample.

Procedure. The 148 pupils who had been tested on the three occasions of the project were divided into four categories: primary, intermediate, white and nonwhite. The smallest cell, primary white, contained 21 pupils. Twenty-one pupils were randomly selected from each of the remaining three cells. A repeated measures analysis of variance was then done on the data.

Results. Other than the known age difference among groups, their characteristics were essentially the same (see Table 6). A repeated measures analysis of variance was done. The significant effects noted were predictable ones and have shown up repeatedly in the data from the project. That is, in the class effect, the intermediates were achieving at a higher level than the primary. The other significant main effect of time simply noted that significant achievement occurred during the three semesters. The significant class x time interaction underlines the often noted finding that the primary children achieved at a faster rate over the three semesters than did the older children in the intermediate classes.

Race as either a main effect or in interaction with other effects was insignificant. Race, therefore, could not be considered a differentiating factor in the achievement of our EMR sample.

Discussion

As previously noted, the findings reported in this paper are admittedly post hoc and should be interpreted with caution. The Wide Range Achievement Test has certain drawbacks as a measuring device, and the passing or failing of one or two items can make a significant difference in the resulting grade level. Despite these cautions, however, the directionality of the results is of some importance. The data supporting the refutation of some cherished beliefs are rather tenuous. The fact that clearly positive support for the assumptions was not forthcoming, however, should at least cause educators to examine the assumptions critically rather than accept them on the basis of face validity. Explanations for the results are conjecture at best, but the findings of the research ap-

TABLE 6
Characteristics of EMR Pupils as a Function of Race and Age

	N	$\bar{X}$ C.A.	$\bar{X}$ I.Q.	% Male	% Female
Primary White	21	98.4	67.9	66.7	33.3
Primary Non-White	21	98.8	65.1	71.4	28.6
Intermediate White	21	133.8	70.0	52.4	47.6
Intermediate Non-White	21	143.6	65.9	52.4	47.6

pear to conflict with some well accepted beliefs of educators.

The number of teachers who have said the equivalent of, "It takes me until October to get them to where they were last June," may well be legion. Yet, at the end of the summer vacation, the data showed no significant loss of skills in any of the three academic areas measured. On the contrary, reading ability increased, particularly among those at the primary level. The results are difficult to explain, particularly in view of the rather widespread feeling that academic skills decline during the summer months. It may be that the abilities measured by the Wide Range Achievement Test are not those considered by teachers when they comment on a lack of retention. Specific skills in the basic subjects may be retained, but the ability to organize knowledge and the memory for information concerning a variety of subjects may diminish. Or, what teachers may really be reacting to is the lack of discipline and loss of work habits brought on by the 3 month absence from school. In any event, the lack of a measurable loss casts some doubt on a traditional assumption and further research designs should have the teachers specify the areas in which they note the disparity. This is particularly true in terms of the results of the reading tests. An increase in skills during the summer vacation is in the opposite direction of the hypothesized change. Explanations for this change must await future research. It is possible that children use the summer vacation period to consolidate skills they have learned and to apply them in relation to their environment. At the primary level the reading skills are minimal, but they do enable a child to understand signs and advertisements on television. These skills are used every day. Given an appropriate beginning, a child may do much to improve his skills to the degree necessary to function in his environment. This may explain why the intermediate group has a smaller increase. This group is already approaching a point where it can cope with most of the reading necessary to function among its peers. In both groups, arithmetic and spelling are of lesser importance and receive a minimum of environmental stimulation. In the population involved, arithmetic skills beyond the handling of money have small functional value.

The second assumption, that summer school has a positive effect on learning, appears self evident. In many locations a remedial summer school class is the means whereby the reluctant scholar passes from one grade to the next. Summer school has been a part of the educational scene since the beginning despite the fact that there appears to be little research that attests to its utility or lack thereof. Still, it is disturbing to note that attendance at the one-month academic summer school resulted in no measurable academic gain in reading and arithmetic. The contrast between these findings and those of students who did not attend summer school casts strong doubt as to whether a traditional summer program is educationally worthwhile. It may be, as previously stated, that the child needs the summer to consolidate previous learning and pressure to extend performance beyond his ability to assimilate is to little avail. In view of sporadic pressures to extend the school year to 11 or 12 months, further research appears in order.

The third assumption, that changing a child's teacher has an adverse effect on academic learning, is given less credence than the first two. Nevertheless, children are usually considered susceptible to a change in routine. This is a problem that besets families in the armed forces and others leading a transient existence. The findings that a change of teachers made no significant difference in achievement are interesting, but they must be evaluated in the context of the information available. The children involved did change teachers, but some may have gone on to better instruction while the reverse applied to others. It appears plausible to hypothesize that the positive and negative aspects equalized each other and the results were not significant for the group. This, however, does not really answer the question regarding the changing of schools. These children changed teachers but there was usually a minimal absence from the classroom. What is probably more important, the children stayed in the same school system. Each of the assumptions mentioned has some basis, whether this basis is logic or experience. It is quite possible that teachers have noted problems in children's moving from one system to another. These problems would be magnified in the case of children who have moved many times without ever having time to become adjusted to the idiosyncrasies of a particular system or teacher. For the purposes of this paper, however, it is apparent that changing teachers and changing schools within the school system had little effect on the level of performance of the children involved.

The fourth assumption could be phrased as a true-false question. It would be a rare educator who would not agree with the fourth assumption. If teaching is the critical element of our educational system and teaching is what takes place in the classroom, it goes without saying that those who do not avail themselves of the teacher's time by reason of absenteeism will learn less. The child who attends school will learn more than the one who does not go to school at all.

Among those tested, differences in achievement level in reading and spelling were noted when the high and low

John J. Cook *is Coordinator, Research Design and Administration, Division for Handicapped Children, Department of Public Instruction, State of Wisconsin, and* Romaine M. [illegible] *is Program Administrator, University of Wisconsin Center on Mental Retardation, Madison.*

This research was done as a part of a larger study which was supported by a contract (OEC-[illegible]) from the US Office of Education. Use of the University of Wisconsin Computing Center was made possible through support from the National Science Foundation and the Wisconsin Alumni Research Foundation (WARF), through the University of Wisconsin Research Committee. The study was conducted at the Division for Handicapped Children, Wisconsin State Department of Public Instruction.

absentee groups were compared. Over the three semesters of the project, however, both groups grew academically at a comparable pace. That is to say, the low absentee group was ahead at the beginning of the study, and by neither gaining nor losing ground remained in the lead. It is possible that academic growth can be likened to physical development and that the twin studies concerning motor training are highly relevant. It may be that the schools are merely teaching the basic knowledges and skills a little early and that the nonattending child learns the basic skills at his own pace. Due to the complexity of learning above the elementary grades, the high absentee group could be expected to reach a peak within the next year or two, while those attending more regularly will continue to improve.

In summary, data was available which permitted a tentative evaluation of four rather common educational assumptions. The fact that the subjects were categorized as educable mentally retarded and came from a disadvantaged area in the central city tends to diminish the possibility of generalizing from the findings. Nevertheless, the results are provocative and it would be well if specific research projects were designed around these and other assumptions which appear to have a minimum of factual support.

THE ATTACHED ARTICLE HAS BEEN ACCEPTED FOR PUBLICATION BY THE JOURNAL OF EDUCATION RESEARCH AS OF DECEMBER, 1975. THE NAME HAS BEEN CHANGED TO TEACHER EXPECTANCY AS IT RELATES TO THE ACADEMIC ACHIEVEMENT OF EMR STUDENTS. CERTAIN CHANGES HAVE BEEN REQUESTED BY THE EDITORS OF THE JOURNAL BUT HAVE NOT BEEN INCLUDED IN THE ATTACHED ARTICLE. INTERESTED READERS MAY REFER TO THE PUBLISHED ARTICLE IN THE JOURNAL INDICATED ABOVE.

Abstract

In view of the current interest in the role teacher expectations play in the education of children, teachers in charge of 18 special classes for EMR children were asked to indicate their expectations for each of their students for the year ahead. Tests administered at the beginning and end of a three semester period disclosed no statistical relationship between expectations and academic achievement.

TEACHER EXPECTANCY AND STUDENT ACHIEVEMENT

ROBERT H. SCHWARZ AND JOHN J. COOK

The education of children has long been subject to the whims and vagaries of specific teachers, principals, school boards and budgets. This is particularly true in reference to educational programs focused on those considered culturally deprived and/or marginally retarded. Professionals in the field are aware of the negative results of research concerning a number of early childhood education programs as well as the findings of numerous studies questioning the effectiveness of current efforts in special education. Parents are beginning to doubt the omnipotence of the school and the business of educating children is experiencing a turbulence that was unknown until recent years.

Under these circumstances, the eager and ofttimes uncritical acceptance of the dramatic results reported in *Pygmalion in the Classroom* (Rosenthal and Jacobson, 1968) comes as no surprise. The idea that the teacher's expectation is a critical variable in the success or failure of a child solves many of the problems raised by negative findings of research studies. Parents are ever ready to grasp at the suggestion that their child could do better. The opportunity to free themselves and the child of the responsibility for poor performance, while attributing it to a teacher who simply did not expect enough, bolsters egos bruised by a lack of academic achievement.

The wide publicity accorded Rosenthal's results (*Time*, September 20, 1968; *Saturday Review*, October 19, 1968) and the positive statements by a noted figure in education, Robert Hutchins (San Francisco Chronicle, August 11, 1968) have done much to turn the findings of a research study into an accepted fact.

According to Rosenthal, the information that a child or children can be expected to "bloom" serves to change the teacher's perception of the student's potential. The teacher believes the child can do better, the child responds with accelerated performance or intellectual growth, and the self-fulfilling prophecy is realized.

The acceptance of this phenomenon is particularly interesting in view of the mixed success reported (no changes in grades 3-6) and the critical reviews of Rosenthal's book by Thorndike (1968) and Snow (1969). The reviews suggested that although there may be some validity to the relationship between teacher expectancy and pupil achievement, the data reported in Pygmalion in the Classroom was so poorly developed and misleading that the research was of little or no value. In an article in Psychology Today, Rosenthal (1968) cites studies of examiner bias supporting the concept of a self-fulfilling prophecy. These findings are augmented by a 1969 dissertation by Haskett. Claiborn (1969), however, was unsuccessful in replicating Rosenthal's (1968) study.

In considering the concept of teacher expectation, it appears that a more realistic measure could be obtained by asking the teacher to estimate the future progress of each child in her classroom. In this case she is free of the opinion of an authority figure and can respond in terms of her knowledge of and aspirations for each child. Under Rosenthal's conditions, there are expectations for the teacher along with the inherent implication that a lack of success by the designated "spurters" can only be attributed to poor teaching techniques.

The idea that children perform in terms of what teachers expected of them is very attractive and has wide ranging implications for education. Thus far, however, the data have been less than conclusive. As a means of augmenting the limited information available, the present study was designed to determine if teachers' expectations for their students were related to the academic progress of the children.

Method

A project was conducted over a three semester period using classes predominately from the inner core area of Milwaukee. The children involved ranged from 6-0 to 12-8 in age with the mean being 9-8. All were enrolled in special education classes for the educable retarded. The IQ's ranged from 48 to 82 with a mean of 68 and an SD of 6.9. These scores were obtained by school psychologists using the Binet or the WISC. In the attempt to determine the pervasiveness of the expectancy phenomenon, teachers were asked to rate the individual pupils in their class with respect to academics (subject matter skills), social behavior (class adjustment) and language (self-~~expression~~ [illegible], written and verbal). Only the results of the analysis of academic expectancy are reported in this paper.

The instructions given the teachers were:

On the basis of your knowledge of the above student, please indicate the extent to which you feel that he (she) will show marked improvement during the coming year in the indicated areas of functioning. That is, each child should be rated in each area from 0 (no change or marked improvement) to 100 (certainty that he will show marked improvement). Consider each child separately and put a check mark above that number of the continuum which reflects your estimate of the likelihood that the child will show marked improvement.

Academics (subject matter skills):

0 10 20 30 40 50 60 70 80 90 100

Eighteen teachers responded to the request to fill out the forms. Definitions of terms in the instructions were clarified during an in-service session. It was stipulated that the child being rated should have been in their class for at least two months prior to their initial attempt to indicate their expectancy of his potential. Wide Range Achievement Testing (WRAT) scores had been obtained from the pupils at three points in time during the project. The resulting scores were used to index the actual academic achievement of the pupils. Only times one (at the beginning) and three (at the end) are considered in this study. Of the 220 pupils who were both rated and tested at the start of the project, 136 were available for testing at the termination of the project and constituted the sample for the present study.

Results

The product moment intercorrelations shown in Table 1 were

Insert Table 1 about here

obtained on a number of salient variables which could be related to the academic expectancy (AE) score. It can be noted that expectancy did not bear a significant (r of .17 for N of 136 required at the .05 level) relationship to the gain scores in reading, spelling or arithmetic. However, the AE scores were significantly related to the scores obtained at times one and three, which in turn were highly intercorrelated as might be expected. CA and IQ had zero order correlations with the AE scores. In passing, it is of interest to note the negative and strong relationship between CA and the three gain scores. This finding has appeared consistently in the data from this project. It underlines the often observed phenomenon in EMR pupils of a slowing down or asymptotic effect to the learning curve with respect to academic subject matter.

Since our approach to obtaining an index of expectancy was somewhat different than those methods used by other workers, it was considered informative to ascertain those factors which could account for the variance on the expectancy scores. A stepwise regression analyses was done with the AE scores predicted by CA, IQ, reading, spelling and arithmetic scores obtained about the same time as were the AE scores and 16 vectors associated with the class membership of the pupils. Table 2 shows the results of this analysis.

Insert Table 2 about here

Some 50 percent of the variance of the AE scores accounted for by the variables entered in the predictor equation. The combined class membership vectors accounted for the largest part, some 41 percent. The other significant contributor was that of arithmetic, about 8 percent. Reading, spelling, CA and IQ were insignificant contributors to the variance of the AE scores.

An analysis of the composition of the AE scores disclosed that as a group the teachers were less than optimistic regarding their prognosis for marked academic improvement on the part of their pupils. The mean rating was found to be 42 percent. Further, teachers varied widely in their expectancies. Some were found to be routinely high (highest average =

67 percent) while others were low (lowest average = 18 percent). In all, expectancy scores ranged from 10 to 100. The zero order correlation between Ch and the AE score in Table 1 is reflected in a consideration of the AE scores for the primary and intermediate classes. The primary classes' AE scores ranged from 14.4 percent to 66.3 percent with an average of 41.8 percent, while the intermediate range was 15.0 percent to 67.5 percent with an average of 42.0 percent.

Discussion

As is apparent from Table 1, the relationship between the expectancy of the teacher and actual academic progress by the pupil is not significant. The fact that teachers are not always accurate in the appraisal of their students is not new. A study by Pegnato and Birch (1959) found that less than one-half of the gifted children in a classroom could be identified by the teacher and she was likely to designate a substantial percentage as gifted who did not have the intellectual ability to qualify.

Of particular relevance is the question raised regarding the concept of a self-fulfilling prophecy. The results are clearly not in accord with the concept that the child will perform in terms of the teacher's expectations. A review of the rating sheets indicates that high hopes were held for a number of children who gained almost nothing during the three-semester period. Conversely, a number of children for whom minimal expectations were held did quite well.

It could be said that the inability of the teachers to appraise the potential of the children changes the context from that of Rosenthal, who told the teachers which children would "bloom." Such, however, is not the case. The accuracy of appraisal is only incidental. The critical element in the self-fulfilling prophecy concept is that the teacher believes the child will do well and this and this alone makes the difference. In this study, teachers stated their beliefs and no relationship could be found.

The conflicting data that has been generated by publicity given the idea of a self-fulfilling prophecy in education makes it difficult to take an unequivocal position. While Rosenthal and Jacobson (1968) based their research on intelligence test score gains, the data reported above resulted from an achievement test. While they report positive findings between expectancy and intellectual growth, no relationship between expectancy and achievement could be determined in this study. Thus, until such time as adequate research sheds much more light on the impact of teachers expectancy, premature publicity, particularly in the popular press, cannot be justified.

Significant portions of the variance of the AE scores obtained in this study were accounted for by two predictor variables entered into the regression equation -- class membership and arithmetic scores measured at the time the AE scores were obtained. Class membership accounted for 41 percent of the AE score variance and simply means that if one knows the particular class a pupil is in, then he can predict fairly closely the AE score the child will get. The finding that the AE score was unrelated to CA (or the primary -- intermediate breakout of the classes) may help to explain the point made in the first paragraph of the discussion section. That is, teachers apparently are unable to accurately differentiate the achievement potential of their students because they seem to use the wrong information in arriving at their estimates. With this population, CA consistently bore a negative relation to rate of achievement. Yet this fact was overlooked by the teachers in arriving at their AE scores. Since the reliability of the AE scores was not obtained, it is difficult to say exactly what determines the achievement expectancy that a teacher has for her pupils. Arithmetic achievement scores explained 8 percent of the AE score variance which leaves 92 percent unexplained, except for the 41 percent which appears to reside in unique characteristics of the teacher at the time of estimating the AE score.

REFERENCES

Claiborn, W. Expectancy effects in the classroom: A failure to replicate. _Journal of Educational Psychology_, 1969, 60(5), 377-383.

Haskett, S. An investigation of the relationship between teacher expectancy and pupil achievement in the special education class (doctoral dissertation, University of Wisconsin, 1968).

Pegnato, C. and Birch, J. Locating gifted children in junior high schools. _Exceptional Children_, 1959, XXV, 300-304.

Rosenthal, R. Self-fulfilling prophecy. _Psychology Today_, 1968, 2(4), 44-51.

Rosenthal, R. and Jacobson, L. _Pygmalion in the Classroom_. New York: Holt, Rinehart and Winston, 1968.

Snow, R. Unfinished Pygmalion, a review of Pygmalion in the Classroom in _Contemporary Psychology_, 1969, XIV (4), 197-199.

Thorndike, R. Pygmalion in the classroom: A review in _American Educational Research Journal_, 1968, 5(4), 708-711.

Table 1

Intercorrelation Table of Variables Discussed in Study (N = 136)

	1	2	3	4	5	6	7	8	9	10	11	12
1. CA	--											
2. IQ	.02	--										
3. Acad. Exp.	.04	.06	--									
4. Read. 1	.34	.03	.23	--								
5. Spell. 1	.38	.14	.21	.86	--							
6. Arith. 1	.70	.27	.16	.73	.60	--						
7. Read. 2	.31	.01	.21	.88	.74	.60	--					
8. Spell. 2	.36	.08	.30	.86	.89	.45	.88	--				
9. Arith. 2	.61	.27	.22	.67	.66	.86	.60	.62	--			
10. Read. Gain	-.32	-.04	13	-.35	-.33	-.34	.14	-.06	-.23	--		
11. Spell. Gain	-.31	-.12	.09	-.30	-.31	-.30	.03	.11	-.20	.67	--	
12. Arith. Gain	-.37	-.08	.04	-.31	-.25	-.33	-.18	-.23	-.03	.30	.06	--

Table 2

Proportion of Variance of Academic Expectancy Scores Accounted for by Relevant Variables

Source of Variance	Amount of Variance Accounted for	Sig. Level
Class Membership	40.03%	<.001
Arithmetic	7.40%	<.001
Reading	1.19%	.090
Spelling	0.03%	.778
CA	0.02%	.861
IQ	0.00%	.936
Unaccounted for	52.33%	

Reprinted from **Bureau Memorandum**, Volume IX, Number 4 (June 1968).

PARAPROFESSIONALS IN THE SCHOOL - I

The following paper is the first of a series of articles on the paraprofessional's role and function in the school. Initially, most of the articles will be based on talks given at an Institute held in Madison August 28, 29 and 30, 1967. The basic intent of the Institute was to enlarge the repertoire of the teacher of the educable mentally retarded in her use of teacher aides in the classroom. Both the teachers and the aides are participants in a research project investigating the use of teacher aides in relation to the size of the class in which they are used. A concern in the project up to this point had been to prepare the aides to work with teachers. It became apparent, however, that it was as important, if not more so, to prepare the teachers to work with aides.

It had been determined that the aides in the project were, as a group, performing all the functions which could be assigned to that role within the existing educational framework. It was therefore reasoned that the goal of the Institute could be achieved by providing the opportunity for within group communication with respect to the role and function of the aide as carried out within the context of the project.

THE ROLE OF THE TEACHER AIDE IN PUPIL MANAGEMENT — Sister Mary Sheila, St. Coletta's School for the Retarded

My first years of teaching were in the back hills of Kentucky. I was a so-called high school teacher in a completely rural area. One day a group of the local gentry were gathered around. They were discussing some of the problems that they were having and one of the men made mention of a mule. Another man said, "Mule? I thought you had a tractor." The first man said, "Well, yes I do, but you got to feed the goldarn mule anyhow, so that's what I use." I think very often in our own lives we are very much like this farmer. We keep clinging to the goldarn mule just because he is comfortable and familiar. We are a little bit hesitant to go ahead and try a high-powered tractor for fear it won't be quite what we thought it would be, or that it might run away with us.

Many times in our lives as teachers we have wished that we could bilocate or cut ourselves in two in order to be all over the classroom at the same time trying to help everybody. This is the way it always is, everybody wants you at the same time. Over the years more and more of these educational tractors have become available, and yet, we as teachers in general have pretty well clung to the old mule. Today I would like to talk to you about probably the highest powered tractor that is presently available to us as teachers–the teacher aide.

Effective use of Teacher Aides

In order to use a teacher aide efficiently it is necessary that the teacher first have a very clear vision of her own role as a teacher. She needs to understand the hierarchy of her own instructional tasks, and then decide which ones fully utilize her skills and training, and which ones have involved her simply as an organizer, an expeditor, or as a director. It is in this latter category where the teacher aide can be the most usefully employed. In this main category there are probably three main areas which should be considered: namely, those of preparing, instructing and evaluating. In each of these areas, the aide can spell the critical difference between pupil control and pupils out of control. I am not talking about behavior, I am talking about productivity.

Those of you who have heard me harp on this same old string before probably know what I am going to say. One of our key problems in special education today is the fact that children are sitting in the classroom hour upon week, upon month, upon year, losing a great deal of their time in school simply because they are devoting their time to educationally unproductive activities.

During the last two summer sessions (I am involved during the summer in our teacher training program) I have had my student teachers check off how much time the youngsters are actively involved in their education, and how much time they are not involved. By not involved I mean that they are sitting there waitingfor help, they are socializing, or sometimes they are sitting there anti-socializing. They are busily destroying instructional materials, they are gazing out the window, they are inventing excuses to leave the room, or doing a great many things that are completely unrelated to the learning process. I would imagine that if I would ask you to hazard a guess, you would probably guess far below what the actual percentage of this lost time is. I once spoke to a very experienced administrator who said he guessed it at about 50% of the time. I was horrified at this. I really couldn't believe that it could be true. But, the older I get myself, the more I think that while he may not be completely accurate, he is probably closer than I would have suspected.

Time Waste in the Classroom

Last year in our summer session, even by the most lenient of standards, we found that our youngsters were wasting a little better than 30% of their time. This summer we did the same project again, and we found out that it was 28% of the time. Remember this is one teacher working with a group of children, and the numbers of the groups range from perhaps ten to sixteen or eighteen youngsters.

I followed this through personally on a little research project I did for one of my classes at the University. I looked at one special education classroom, and found the youngsters wasting (and I guess that's the most accurate term I can use here) 40% of their time. To me this is tragic. I have seen the sad end results of this in the young people who come into Dr. Coletta's for our rehabilitation program. Some of them come to us seventeen, eighteen and nineteen years old, with I.Q.'s in the high educable range e.g. but who are achieving at a functionally illiterate level, far below their grade expectancies. This should be a burning concern to us in special education. Why have these kids sit for six, eight, ten years in a special education classroom, and come out of it as illiterates? Why don't they learn more than they do? One of the reasons for this tremendous percentage of time waste is directly related to our classroom management techniques.

Classroom Management Techniques

Hopefully, the teacher aide could come in here and fulfill a very vital function. When you stop to analyze why it is that your youngsters in the classroom are sitting there doing things that are not related to their education, it is usually for one of three reasons, or related to the three areas mentioned before. Either the teacher is so busy teaching somebody else that she doesn't have the time to go back and help the child a little bit so he can get on with an assignment, or simply because in his independent work time he does not have the materials that were tailored made to meet his learning level, and his learning characteristics. The reason is the teacher just doesn't have the time. She may have between ten and sixteen youngsters to prepare for, and she can't possibly make up all the little devices for every single child for every single day. It can't be done, and so very often then the youngster is again sitting there wasting time.

Then thirdly, you will find the youngsters who simply will not go ahead until they have been reassured that what they are doing is correct. All it takes is one quick cruise around the room to check up on whether or not the child is doing it correctly. But until this has been done by the teacher, often the child will simply sit there. Therefore, I think that anything that a teacher can do to insure the continuity of her instruction and avoid interruptions is going to improve the productivity of her pupils.

Extend Teacher Effectiveness

In the area of planning we have to determine, first of all, what we're going to teach, to whom this will be taught and which materials are going to be most useful. Once these technical level decisions have been made by the trained teacher, we can use the aide to extend the teacher's effectiveness in preparing the needed materials. Again, these would probably have to be according to teacher specifications. For example, take as simple an idea a thing as the preparation of flash cards. I have found in working with primary level youngsters that it helps to give them their own individual flash cards for reading, for spelling and for number work. It takes a lot of time to do this sort of thing. Once the aide knows what to do, this is one way of helping the teacher extend herself. Currently, the preparation of transparencies is becoming a more important thing in our lives. It can be very time consuming and is something that with a few minutes of instruction can be done by the aide. And then, too, there is the devising of instructional games. All of you have had the experience of working out some little way of playing bingo, or playing old maid with the youngsters. The only problems are first - you have to have time to get it ready, and secondly - you have to have time to teach the children how to play. If it weren't for these two things, the youngsters could continue in this sort of an activity almost indefinitely, and all the while they were doing it they would be learning. So here again, you can see where the aide could be very useful. The teacher gets the idea, passes it on to the aide, who prepares the materials and when the youngsters are on independent work time, the aide herself can sit down and teach the youngsters to play the game. Later on she can withdraw and the group can go on independently.

The making of charts is another area where very often we have taken a lot of unnecessary teacher time. Once the main idea is there, I think it is very useful to have the aide develop the chart.

One of the techniques that we have found at Dr. Coletta's that is very useful is the development of a file of mimeograph masters for each teacher. How wonderful it would be to have an aide who could type up these things, and develop this file for you in a shorter time than the eight years it took me.

Source Material

Another very important aspect of our planning is that of tracking down source material. This is an area where an aide can be very useful. They can review catalogues for the film listings, and film strips that are available down at the library and look for records. Very often the initial rough screening of films, records, film strips and so on can very nicely be done by a lay aide. Later on the teacher is going to want to preview all these things herself, but an initial screening can narrow the number of materials to be previewed by the teacher.

It is a very useful thing to have your lay aide make a listing of all the books that you have in your classroom library. You might be amazed to find the wealth you have right under your very nose. You haven't had time to look for it and once you know it is there, you will be much more likely to use the materials in terms of your own reference, and in terms of what you could have the youngsters do in reference assignments.

Classroom Party

The idea of the classroom party can be used more effectively as a teaching device by using the aide to escort the committees out to purchase the food or other party supplies, or to act as a referee at little committee meetings that the youngsters might have. Again, this sort of thing doesn't take a great deal of teacher training and often an aide can do this quite easily for us.

Many, many resources in our own community simply are never tapped because we don't have time to call up the people and make the contact, and we don't have time to go out and visit the site ourselves to see exactly what it is. This is a good area where the lay aide can come in and make the contacts for us. Go out, run through this factory or run through the park and see what is really there, and come back and say, "Well, you'll have to watch out because there is a bad traffic hazard here," or "I think the youngsters would enjoy this part of the plant, but probably not the last." Again, this is the thing the teacher ought to do before she takes her class anywhere, but life being what it is, I don't think that many of us are really able to do this.

Library Aide

One thing that we have found very useful at St. Colette's is the use of aides in the library. The teacher sends her classroom aide into the library, and then the youngsters are able to go in there during school hours, draw out books and have an adult help them in the selection of books. We have found that this helps the teacher to plan more creative types of assignments. If you know there will be somebody in the library to help the youngsters, you can tell the youngsters, "All right, now you go and find out what kind of leaf this really is." "Go into the library and see if you can find a book that has this picture." If there isn't an adult there, this would be a pointless kind of thing; but with an adult on hand who can give the child some assistance, this can be a real learning situation for a child: one which is not possible unless we have this extra aide helping us in the room.

Understanding Teacher Goals

In all of these activities, though, it is very important that the aide understand what the teacher's goals are. Some of the real problems in working with aides are that they think we are doing one thing, and we think we are doing something else. We can imagine what the child thinks. Perhaps one of the real fringe benefits of having an aide in the room is that it makes those of us who are teachers sit down and think: exactly why am I doing this, precisely what concept am I trying to develop, specifically what skill is involved there? The lay aide can serve as a challenge which will stimulate us to put our goals into a little sharper focus.

Time Saving

When we consider the area of actual instructing, we will find that the aide makes it possible very often to utilize some hitherto completely wasted teaching situations. Sometimes we do things simply from the viewpoint of classroom management. I know I've done this myself. "Look, I've only got an hour and a half for this art project. I'm not going to have the kids spending forty-five minutes of an hour art period getting the finger paint out. So, when they come into the room everything is there. Then we have our neat lesson, and I'm in control. I know what is going on; and when it is finished, the children go sit down and when they are out of the room, I clean up." You see what I've done is cut off at least 50% of the learning experience here simply because I need to manage my classroom. By having an aide in the room many situations of this nature can better be utilized. Many of the ideas could be so much better implemented in our rooms if we could use an aide in this sort of thing. Those of you who have visited some of the Montessori rooms have been impressed with the amount of teaching that takes place over carrying a tray of glasses, or a pan of water, or sweeping up one corner of the room. Many of these activities are a tremendous learning experience but are horribly time consuming. Also very often we have youngsters on several levels, and while what you're doing may be appropriate for one level, it is not appropriate for the others. So very often what we do is just let these activities go right out the window. If we had an aide who could work with a small group who are ready for a certain type of an activity, we would be much less likely to waste the opportunities.

Independent Work Time

Probably the greatest contribution that the aide can make is that of helping the youngsters fully utilize their independent work time. When you stop and think of it, if the youngsters are in the classroom five or six hours or four hours a day, during how much of that time are they actively working with you? Usually it is a relatively small percentage simply because of our grouping for instruction procedures.

One of the greatest handicaps that we find among our older retarded children is the fact that they are lacking in study habits and work skills. These are the things that need to be taught and taught specifically to the children. However, it is rather hard to be teaching them how to study when you are doing active instruction in another corner of the room. I can see a tremendous function for the lay aide's actually teaching the youngsters how to start a job, how to see it through, how to do a little self-evaluating, and then how to make the transfer on to another job. Youngsters are very often unable to make the transfer from one task to another. They are very unable to work on any self-initiated task, and a great deal of learning can take place on these self-initiated tasks. I remember once having a group of youngsters who couldn't learn Arithmetic until I hit upon a motivating device that made them suddenly want very much to learn Arithmetic. I gave them some little self-study books. It was amazing then how well they were able to study those facts and learn them before the deadline came. They were sitting in line waiting to get into the showers studying these number facts. They were walking back and forth to school studying their number facts. They were fighting with each other on the steps because somebody took away their book of number facts. I couldn't achieve this on an instructional interaction, but they were motivated enough to do this independently. That year those youngsters really did show a great deal of achievement gain. This is one area where we as teachers need to take a much closer look, and again, this is one place the aide can help a great deal. The aide is not responsible at all for any of the live instruction. She is floating around the room working with those kids who are studying and doing independent work. Aides can do a great deal of this sort of thing that teachers simply cannot do.

Spelling Tests

Another thing that we waste a lot of time on is dictating spelling tests and dictating reading tests. It doesn't take a great deal of training to say a word, use it in a sentence and repeat the word and wait until the youngsters writes it down. It does take a great deal more skill to develop the phonetic and the structural analysis that is entailed here. This is a teacher's job. I can see the aide very nicely taking those kids over in the corner, e.g. challenging them, and then going ahead and dictating the lesson. I think the same thing applies to reading tests we give. We know as teachers why we give them. They don't tell us a thing about how the children read, but they do give the youngster a chance to get 100%.

Ideas on Techniques

Very often these activities will be successful only if the teacher will give the aide some ideas of the techniques to use. Otherwise all of us tend to use the techniques that were used on us when we were at about this grade level. This is what your untrained aide is going to do because this is what we ourselves would do in a similar situation. If we don't give them some ideas on the techniques to use, they will simply revert back to the techniques of the 30's, 40's when they themselves were in school.

Very often we will find that we waste a great deal of time explaining an assignment to a child. We understand exactly what we want the child to do, but we have to be very specific in how we explain it. This takes us maybe five minutes of explanation while the other youngsters are sitting there waiting for us to come to them. Again, this is a complete loss of time. A lay aide could probably do this much faster. You could tell the aide in about forty-five seconds what it would take you approximately five minutes to tell the children. Tell the aide and let the aide tell the child. Save your breath for something that is perhaps a little more demanding of your professional competence.

Some of the problems that we have run into and somehow circumvented as teachers in communicating with retarded children are going to need to be taught to our aides. To put it in the simplest terminology: "You show them - you don't tell them." Very often we forget about this when we turn the job over to somebody else. We forget how many years it took us to figure the procedure out. What you have to do at the beginning when you start working with your aide is say, "Well look, you explain it this way, and you show them this way."

You might think, "Well then what time have I saved? First I do it for the aide and then the aide does it for the child. I have used the same amount of time." Surely this is true, but it is an investment. It takes you a long time the first time you tell the aide, and a little less time the second time, and a little less than that the third time. After that the aide takes over and you can move out into fields that require a little more technical competence.

Enrichment Activities

We pay a great deal of lip service in Special Education to the provision of the so-called enrichment activities. We all know they're terribly important, but deep in our hearts we also know they are dispensable. Because of the harsh realities of life, we very often simply dispense with them. An aide, however, could at least make some of these dreams come true. For example, the daily read-aloud period. We know that these youngsters from the time they come to us until they leave us really should be read aloud to every day. The lay aide can take the youngsters off during some of their independent work time and read to them.

The teacher should select the material that will be read. She is the one who is thinking of the correlations between what is being read during story period, social studies or science, or reading; but the aide can do the actual time consuming part of the task.

This is also very true in terms of oral speech. Very often we are just so busy that we never take time to talk to the children about what they are doing. Oral speech is one of the really weak areas with our youngsters. They simply do not have the verbal mediators with which to think. This is another area where we can use an aide a great deal. Teach the aide to do just as the parent would do in the home with a smaller child. "Oh! you're doing this, and then talk about it." Use the terms, use it in sentence form, give these youngsters many experiences of talking about, of verbalizing about what they are doing. I think if we would do this on a planned basis over a number of years, we would find that many of the language deficits that our youngsters have would be to a great extent remediated. This is a very simple technique. It would be a wonderful thing if we could have time for it in the classroom. Most of the time we haven't, but an aide could do this very, very well.

I think the same holds true of the problem of giving the youngsters music appreciation or art appreciation. These are all things that we do believe very important for children, but if it is a choice between, do I teach them arithmetic, or do I teach e.g. them music appreciation? He has to know his arithmetic and so the music appreciation goes down the drain. Have the aide, for example, play "Peter and the Wolf" until the youngsters are able to identify the various musical instruments or the theme of some of our better works of music. The same holds true of Art. If we can do anything to broaden the horizons of these youngsters, we have done just that much better a job. If we can't do it ourselves, it is our responsibility to find somebody who can.

Positive Reinforcement

When we come to the area of evaluation, we have to realize the very great importance of giving the youngsters immediate feedback on whatever they do. The young one does his arithmetic, he puts the workbook on your desk and when everybody is gone that night, you sit down and put the marks on it. The next morning you hand it back to the child, and he couldn't care less. He has forgotten what it was all about yesterday. Very often you have to reteach the whole thing before the youngster is able to really benefit by the corrections that you have given. Here you can see the very great value of having an aide who is walking around the room correcting right at the same time the youngsters are working. It is the responsibility of the teacher to analyze the errors that youngsters are making. This is the feedback she needs in order to adjust her own instruction. I don't think anybody would quarrel at that point. On the other hand, sitting there with mountains of papers and workbooks at the end of every day and putting on the red checks or stamping on the gold stars over and over and over again is not providing much feedback.

With an extra person in the room it is possible to work on perhaps a more elaborate reinforcement system. As soon as the youngster does something correct, the lay aide is right there to score his marker, or to give him whatever you are giving him as reinforcement. You could certainly accelerate the modification of behavior if you were able to keep up this immediate reinforcement.

The benefits accruing from the presence of the lay aide in the room could be manifold. I think that the actual value is really in proportion to the teacher's ingenuity in utilizing the aide and in her willingness to invest in doing a little development of the aide's skills. I think many of the little teacher tricks, I am not talking about technical skills, that we all pick up in our years of teaching experience can be very baffling to the uninitiated outsider.

Interest of the Aide

It has been our experience in working with our aides at St. Coletta's that you get a higher level of performance and interest with the aides if the teacher shows a willingness to invest a little of her time in teaching the aide some of these very useful skills. The aide is helping you. I think you should show willingness to help her.

Aides are more interested if the assigned tasks involve direct interaction with the children rather than a heavy preponderance of housekeeping jobs. This is one thing that has concerned me in the very few articles that I have seen published on the use of the aide in the classroom. You look at a list of suggestions and maybe out of thirty suggestions, twenty-eight have to do with housekeeping. Paint the doll dishes, make the drapes, wash the dishes, pour the milk, take the child to the toilet, these sorts of things. Wouldn't it be great to get all these things off our backs? Some of them the aide can take care of and is willing to take care of, but e.g. if that's all they ever get to do, as one of our aides once said to me, "Sister, I'm going to have to quit coming, I can do this sort of thing at home." She had a point. I think you will find that the aide who perceives her work as necessary is going to be a much more effective aide. Therefore, the work that she gets to do should not be just the busy work sort of thing, the glorified secretary or the glorified maid or the glorified custodian. While it is true that we would like to have many of these tasks relegated out of our hands, I think we have to show a certain balance in what we give the aide to do. The assignments should follow a certain pattern or routine. If they are definite and allow the aide to use a little initiative and ingenuity, she is going to be a lot more interested. The work of the aide as much as the work of the children must be planned for. Perhaps this does add a little bit to your planning burden; but when you think of it as a long-range investment, I think you'll see that it is one that pays the big dividends.

Use of Aides

In a periodical called the Professional Growth for Teachers, the fourth quarter issue of last year, there was an article on the Use of Teacher Aides which presented two types of programs. I call it to your attention because it might have a few ideas of use for you. The first teacher described how she used an aide - mostly in housekeeping activities. The second teacher brought out the idea that a teacher using a number of aides could run a completely individualized remedial program. If you take a look at the two extremes, you can find a modified version which would fall somewhere in between.

If you are fortunate enough to have as your aide a teacher who quit because she had little children and now is coming back as an assistant, she can do a great deal different type of things than the aide who has always been a housewife and who knows a good many mothering skills but doesn't know a great deal about classroom-type activities.

Evaluation of Aides

The evaluation of the teacher-aide program has been up to this time, at least to my knowledge, pretty much on the anecdote level. This is particularly true in Special Education. A study that was conducted by Cruickshank and Haring summarized it. It rather amused me. I don't know how much grant money they had involved here, but out of it they came up with a summary of three points. First, the teacher said they had new and more materials with which to teach. Secondly, they planned better instructional programs for the children. Thirdly, they gave more individual attention to the children.

Hopefully, at the end of the project in which you lucky people are all now engaged, we will be having a little more objective evidence about the value of the teacher aide in the classroom. You may turn out to be the super salesmen who are going to sell the educational tractors, the ones that would be of most benefit to our pupils. Maybe you will be the ones who are going to turn some of those tired old mules out to pasture.

THE ATTACHED ARTICLE HAS BEEN ACCEPTED FOR PUBLICATION BY EDUCATION AND TRAINING OF THE MENTALLY RETARDED AS OF DECEMBER, 1970.

THE UTILITY OF MENTAL AGE AS A PREDICTOR OF ACADEMIC ACHIEVEMENT

Robert H. Schwarz and John J. Cook

ABSTRACT

Data resulting from two studies involving 499 mentally retarded children in special classes revealed that the differences between actual and expected achievement tend to increase as a function of age regardless of IQ. While the MA may be a reasonably valid measure of the developmental level of the individual, other factors appear to exert a significant effect on achievement and the MA has questionable ability as a standard for educational placement.

A recent study (Schwarz and Shores, 1969), disclosed a significant relationship between the achievement patterns of a population of EMR students in special classes and their social class as measured on Warner's scale (Warner, Meeker and Eells, 1960). While the findings were generally in accord with other research in this area (Hill and Giammatteo, 1963; Abrahamson, 1952), the differences in achievement were noteworthy in terms of the time spent in an academic setting. In retrospect, the results seemed to indicate that mental age has little or no utility as a measure of academic achievement among culturally disadvantaged retardates. The use of MA as a means of comparing groups has been subject to some criticism (Baumeister, 1967) and discussed at length by Denny (1964). Each felt that comparisons between "normals" and EMR children on the basis of MA alone led to methodological errors.

This, however, does not negate the fact that the MA can be a useful measure in the context of an appropriate research design. Unfortunately, the ease with which it can be computed and the lack of information regarding its validity as a measure of academic potential has led to the use of the MA for placement. If mental age is a legitimate measure of developmental level (see Zigler, 1969), a child should be able to progress in school at a diminished rate in accordance with his calculated state of readiness. As most teachers and administrators are well aware, this is not the case for the majority of EMR children. Few, with an IQ of 75, accomplish three fourths of a years' work during each school year. Or, in terms of mental age, few 16 year old students with an IQ of 75 could compete academically with a normal 12 year old.

However, as might be expected in terms of individual differences, variations in achievement patterns do occur and Itkin (1970), reports the results of a survey of Chicago EMR classes wherein a significant number of the children were found to be non-achievers in relation to others of similar mental ability. Background variables were examined and the non-achievers were found to have less of everything of value to the middle-class culture. Those who were achieving within reasonable expectations tended to come from intact families of semi-skilled workers.

Method and Procedure

The data presented in this study were obtained from both inner-city and rural settings. In part, it resulted from a larger study concerning class size and teacher aides (Blessing and Cook, 1970). These findings were augmented by data gathered in special classes serving a disadvantaged rural population. The children involved had a CA range of 74 to 184 months. IQs ranged from 51 to 82.

Achievement scores were obtained through the administration of the Wide Range Achievement Test (Jastak and Jastak, 1965). Individual results on the reading and arithmetic subtests were averaged and the resulting scores considered to be the level of academic achievement. In each case, this level was compared to the expected achievement in terms of mental age as reflected by the test manual. The IQ of each child was obtained by means of the administration of the Binet or WISC by psychologists certified by the State of Wisconsin.

The subjects were divided in terms of CA and grouped in 10 point increments in IQ. The mean achievement score for each group was determined.

Results and Discussion

The results are as shown in Figure 1. As is apparent from the

****

Insert Figure 1 about here

****

consistent negative slope of all three groups, the culturally disadvantaged, retarded children involved in this study demonstrated a steady decline in terms of expected achievement. The utility of mental age as a variable can be questioned but commonly accepted achievement tests are based on grade level which, in turn, is a function of age. Further, a number of investigators have demonstrated that the learning ability of groups equated on MA is generally similar (Zigler, 1966; Osborn, 1960). Achenbach (1970) carried this concept into intelligence testing and found a high degree of similarity in the response patterns of individuals matched on mental age. Thus, research designed to compare individuals on the basis of mental age seems to produce data which supports the MA as a measure of developmental level. The differences that do occur have been ascribed to motivational factors. Zigler (1969) hypothesized that if equally effective reinforcers were found, retardates and normals of the same MA would function similarly.

Other investigators confirm the idea of differing motivational structures and in a study comparing middle and lower class children, Terril, Durkin and Wiesley (1959) found that middle class children responded best to intangible rewards while lower class children were superior when competing for tangible rewards. In all, research in the school achievement of children indicates that full consideration must be given the child's perception of the school setting. Schools have given little attention to the reinforcement hierarchy of students differing in cultural backgrounds and one result may be the lack of academic achievement found not only in special classes but among students from poverty areas in general.

It is quite possible that the decline in achievement found among the children examined as part of this study was not atypical of many other students attending regular classes in the area. Although the study was structured to provide data which would extend a preliminary evaluation of adolescents (Schwarz, 1969), it may serve as another inditement of public school education for the culturally deprived.

In his book, Hurley (1969) makes the point that a review of poverty programs throughout the country reveals general educational failure. Testing programs in regular classes give evidence of underachievement in relation to ability and school attendance. The system as presently constituted does not seem to be reaching those children who have potential to learn. Even among children deemed intellectually normal the discrepancy between their academic achievement and that of an equal population of more advantaged children increases in relation to the number of years in school.

Robert H. Schwarz is Administrator of the University of Wisconsin Center on Mental Retardation, Madison and John J. Cook is Coordinator, Research Design and Administration, Division for Handicapped Children, Department of Public Instruction, State of Wisconsin.

References

Abrahamson, S. School rewards and social status, Educational Record Bulletin, 1952, 31, 8-15.

Achenbach, T. Comparison of Stanford-Binet performance of non-retarded and retarded persons matched for MA and sex, American Journal of Mental Deficiency, 1970, 74, 488-494.

Baumeister, A. Problems in comparative studies of mental retardates and normals, American Journal of Mental Deficiency, 1967, 71, 869-875.

Blessing, K. and Cook, J. Class Size and Teacher Aides as Factors in the Achievement of the Educable Mentally Retarded, 1970, Final Report to USOE, Grant # OEG-3-6-062620-1879.

Denny, M. R. Research in learning and performance, In Stevens, H. A. and Heber, R. (eds.), Mental Retardation, Chicago: University of Chicago Press, 1964, 100-142.

Hill, E. and Giammatteo, M. Socio-economic status and its relationship to school achievement in the elementary school, Elementary English, 1963, 40, 265-270.

Hurley, R. Poverty and Mental Retardation -- A Causal Relationship, New York: Random House, 1969.

Itkin, W. The non-achieving mentally retarded: Implications of his needs for education and the community, Mental Retardation, 1970, 8, 53-54.

Jastak, J. and Jastak, S. Wide Range Achievement Test, Washington, Del.: Guidance Associates, 1965.

Osborn, W. Association clustering in organic and familial retardates, American Journal of Mental Deficiency, 1960, 65, 351-357.

Schwarz, R. Mental age as it relates to school achievement among educable mentally retarded adolescents, Education and Training of the Mentally Retarded, 1969, 4, 53-56.

Schwarz, R. and Shores, R. The academic achievement of EMR students and social class, American Journal of Mental Deficiency, 1969, 73, 338-340.

Terrell, G., Durkin, K. and Wiesley, M. Social class and the nature of the incentive in discrimination learning, Journal of Abnormal and Social Psychology, 1959, 270-272.

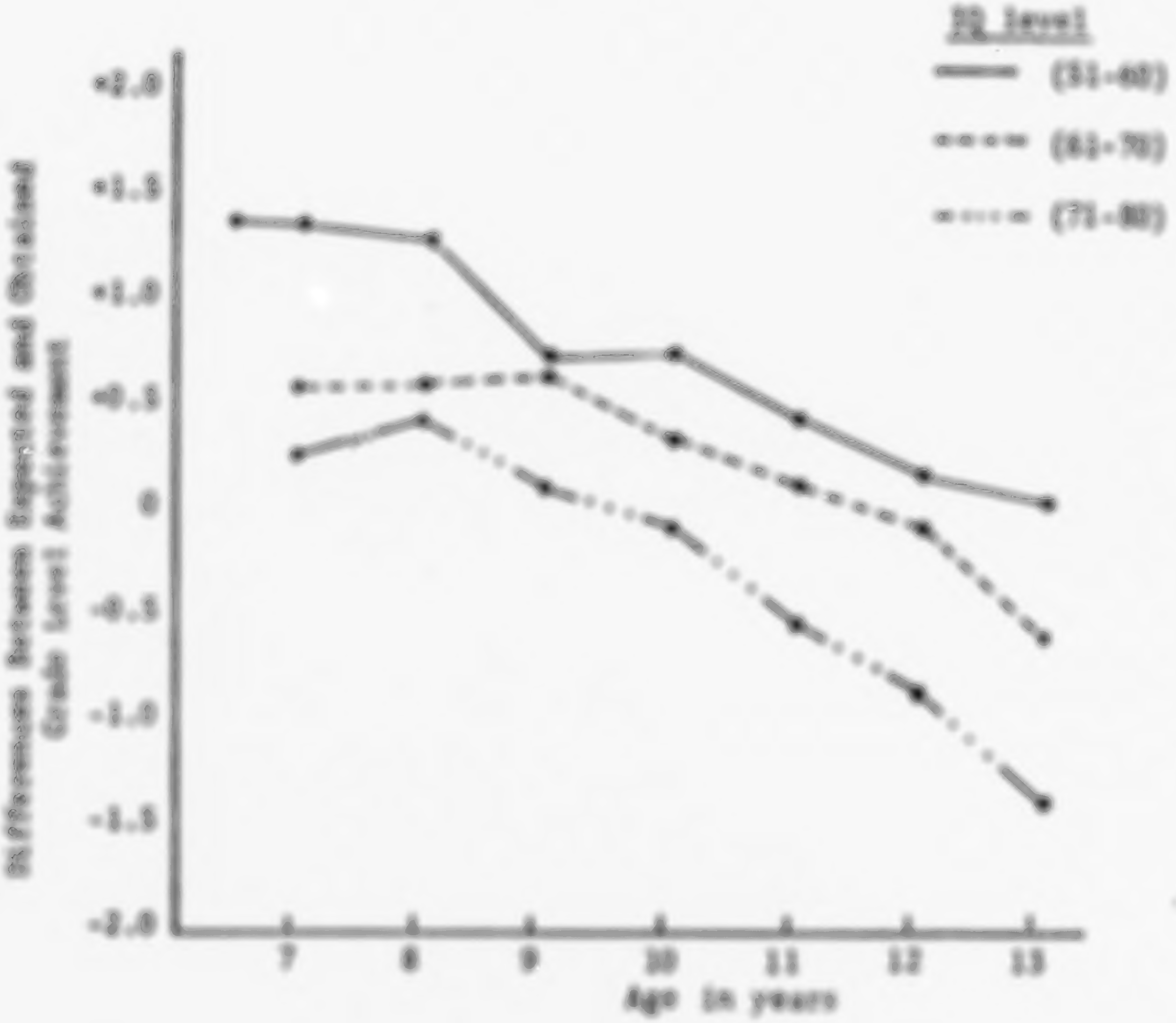
Warner, W. L., Meeker, M. and Eells, K. Social Class in America, New York: Harper and Row, 1960.

Zigler, E. Developmental versus difference theories of mental retardation and the problem of motivation, American Journal of Mental Deficiency, 1969, 73, 536-556.

Zigler, E. Research on personality structure in the retardate. In Ellis, N. R. (ed.), International Review of Research in Mental Retardation, Vol. 1, New York: Academic Press, 1966, 77-108.

FIGURE 1

Differences Between Actual and Expected Achievement
In Terms of IQ and CA



Reprinted from "Bureau" Memorandum, Volume XI, Number 3 (February 1970).

Noise Levels in the Classroom

by John J. Cook, Ph.D.
Coordinator, Research Design and Administration

In the past few years much more attention is being paid to acoustics. We now see schools designed with more open space and with the use of acoustical materials on walls, floors and chairs. It seems to be true that the more open design of a school leads to a better noise level. More attention is being given to the activities which will go on in a room such as music or gym or the use of television and audio-visual equipment. School designers are giving much more attention to the problems of acoustics and the result of this will be a more quiet learning situation for students.

Research findings on the effects of noise on human performance have, in general, been conflicting. Some investigators have reported detrimental effects while others have reported that noise levels present no problem to the human learner. Individual differences in adaptation to and need for control over noise levels probably accounts for much of the lack of agreement on the effects of noise level. It seems safe to say, however, that extremes in noise or in quiet could have a detrimental effect on learning in the classroom.

In the above discussion, noise level has been treated as a cause of variations in behavior and learning. We can, however, turn over the coin and look at noise level as an effect or, in research terminology, a dependent as opposed to an independent variable. Conceived as a dependent measure, we can try to find those factors which influence noise level in the classroom. In a research project just completed, the questions asked were whether or not class size (standard size versus a 50% increase), use of an aide and the age of the children (primary vs. intermediate) had a differential effect on noise levels sampled from the classrooms over the three-semester duration of the project. In this context, we made the assumption that noise level might be a direct and easily obtainable index of classroom control.

The present paper has two purposes: first, to provide descriptive data on noise levels as recorded in the 21 classrooms during the course of the research project. Secondly, we hypothesized that varying class size and use of an aide in two levels of age groups of pupils should be reflected in varying noise levels as an index of classroom control.

Procedure

The ear's response to auditory stimuli is somewhat complicated so that a given increase in level may be important at low intensities but relatively unnoticed at high intensities. As a partial adjustment for this, sound level is designated on a logarithmic scale, the decibel (dB). By definition then, 10 dB represents 10 times as much power as 0 dB; but this is meaningless unless there is agreement as to the definition of zero dB. A close approximation to the maximum threshold of hearing at 1000 cps has been standardized at a sound pressure of .0002 microbars (watts/square centimeter) and is considered zero dB.

Useful tables of sound levels have been established by various governmental and private concerns relating environmental equivalents to the dB scale. Such a table compiled from several sources is shown in Figure 1.

A research project was recently completed in Milwaukee concerned with the effects of variation in use of teacher aides, class size and class-age status on the academic achievement, behavior and language functioning of EMR pupils. During the three semesters of the project, 10-minute audio film samples were obtained at approximately two-month intervals from each of the 21 classes involved in the project. The filming equipment used had the capability of providing noise level readings (Nagra III Recorder through a Sennheiser microphone) and also light readings (Bolex camera and attachments). Thus, with a minimum of effort and expense, we were able to obtain descriptive data on certain environmental factors previously unreported for special education classes.

The schools containing the 21 classes were located in the so-called core of Milwaukee or the near South Side of Milwaukee. The schools ranged in age from two years (Holmes) to 40 or 50 years (Fourth Street and Riley). Noise level did not appear to be a function of a building's age however.

Technique

The microphone was placed at various locations in the rooms to measure the sound levels. The time at each location was no shorter than 10 seconds nor longer than three minutes in order to obtain an adequate sampling of the noise in the room during the 10 minutes of filming. The testing included both the peak noise level or loudest noise observed and the ambient level which is defined as the noise level of the room excluding individually distinguished sources other than continuous operational or street noises.

Results

Figure 1 shows the mean ambient and peak noise levels in dB of the 21 classes over the seven filming periods. Since dB are so relative, a column of environmental equivalents is provided in the left side of the graph. The environmental equivalents have been derived from several sources and are included under the assumption that they have meaning to the reader.

TABLE 1 - Ranges and means of ambient and peak noise levels in 21 classrooms of EMR pupils over seven sampling dates.

Sampling Date	Ambient Range	Ambient Mean	Peak Range	Peak Mean
Feb. 67	55-75	66.6	69-85	75.6
Mar. 67	60-73	65.4	67-87	75.8
May 67	50-70	65.0	63-85	71.3
Oct. 67	48-73	61.1	58-88	71.7
Dec. 67	53-67	57.5	58-78	69.5
Mar. 68	53-75	58.0	67-85	73.2
May 68	38-63	58.2	53-83	71.2

Some agreement has been reached as to what constitutes degrees of noise level in a classroom for normal children. Sister (1968) defined quiet conditions from 40-55 dB, average from 55-75 dB and noisy conditions from 75-100 dB. Since EMR classes typically have much fewer children, it could have been hypothesized that the ambient noise levels in our 21 classes should not extend above the average level.

Figure 1 (and Table 1) shows that the noise levels were well within the so-called normal classroom level, ranging from an ambient mean of 57.5 dB to 66.6 dB. A general trend of decline in noise level over the three semesters can be noted. The reason for this is not at all clear. The peak noise level averaged about 9 dB above the ambient level with some classrooms, however, reaching a peak of 88 dB (see October, 1967, in Table 1) which in terms of environmental equivalents is somewhat louder than a subway train at 20'.

A class (2 levels) x aides (2 levels) x size (2 levels) x time (3 levels) repeated measures analysis of variance was done on the ambient noise level data. A significant class x time interaction (F=4.80, p .03) was noted. Over the three semesters the primary classes were consistently noisier than the intermediate classes with this difference increasing, and by the third semester the intermediate classes were on the average 5 dB quieter than the primary classes. The classes with aides in them were consistently quieter (by about 3 dB) than the classes without aides. This difference, however, was not statistically significant.

A rather complicated second order interaction of aides x size x time was found significant (F 4.14, p .03). This interaction defies explanation by the author but will be reported nevertheless. In those classes with aides, the larger classes maintained about the same noise level while the standard-size class had its noise level reduced and by the third semester an 11 dB difference was evident. On the other hand, in those classes without aides, the noise level remained about the same for the standard-size class and fell for the increased-size class. By the third semester, the difference was 5 dB.

Discussion

It is obvious that the generality of the findings in this report are quite restricted as they were obtained from 21 EMR classes in Milwaukee in what at best could be called a quasi-random basis. Therefore, the discussion must be considered more suggestive than definitive.

Our data have shown that the noise levels in the EMR classes sampled were well within what is considered the "normal" for classrooms. Assuming that most classes fall within this range and given the contradictory evidence of the effects of noise on learning, one is led to the conclusion that the present concern in reducing noise levels in classrooms must have its basis in fiscal and aesthetic matters rather than enhancing learning, per se.

We have also considered noise level as an easily obtainable dependent measure or index of classroom control. Whether this assumption is correct or not has not been supported by any data we have. In any event, it is imperative that we consider unobtrusive measures of this type in the evaluation of the various projects and programs in special education. There are several reasons for this. First, this type of measure is easily obtainable and inexpensive; and secondly, they are non-reactive in the sense of not interfering with or biasing the instructional process as so many of the measures and tests we presently use do.

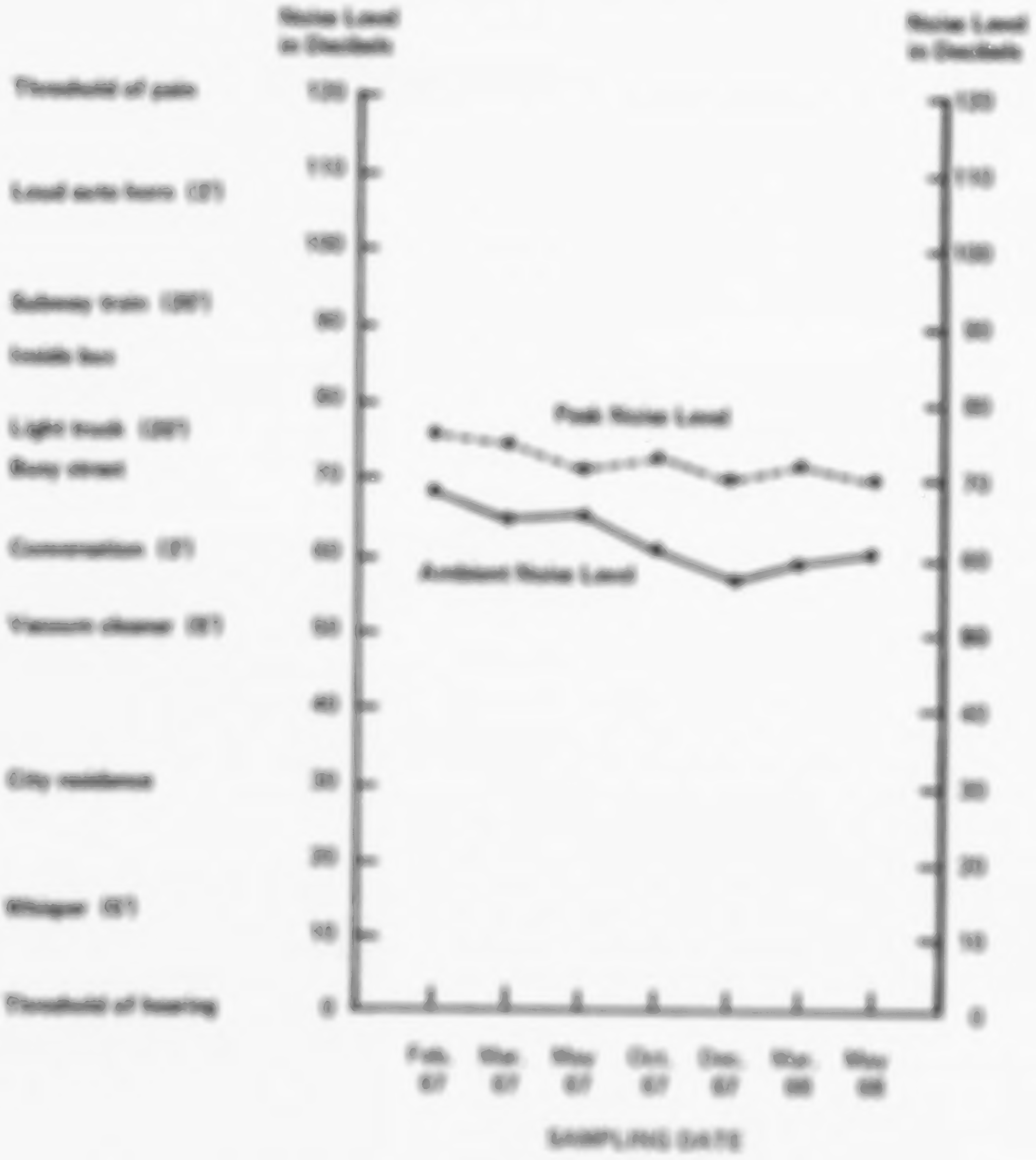
Substantively, our results have shown that primary EMR classes are noisier (more out of control?) than the intermediate classes. The teacher aide appeared to have a marked effect in reducing noise levels (increasing control?) in the standard-size class, but not in the class of increased size.

In general, these findings must be considered very tentative and in need of replication with a larger number of and more representative classes.

References

Slater, B.R. "Effects of Noise on Pupil Performance". J. of Educ. Psychol., Vol. 59, 1968.

FIGURE 1 - Sound equivalents (environmental and at a given distance from noise source) of mean ambient and peak noise levels in 21 classrooms of EMR pupils over three semesters.



Light Levels in the Classroom

John J. Cook

The present paper provides more descriptive data on light levels measured in 23 [illegible] classes over the course of a research project which extended over a three-semester period - spring, 1967, fall, 1967, and spring, 1968. Specific questions asked of the data were:

(1) To what extend did seasonal variations influence light levels in the classroom?

(2) To what extend did the 23 classes meet the state illumination code requirements?

Background

Modern school programs include a great variety of school tasks such as reading, writing, painting, sewing, playing games, and so on, which can impose a burden on the eyes. In each of these tasks adequate light requirements may vary. Schools with divisible teaching spaces must plan special lighting systems that can effectively illuminate both large and small group areas.

Research and observation in classrooms throughout the country indicate no evidence to support the need for more than 30-foot candles, provided optimum conditions are achieved in other respects: such as reflection factor, brightness ratios, glare, etc. It also seems unnecessary to provide 30-foot candles everywhere in the classroom. Rather, a range of lighting might well be provided utilizing a general lighting level of 30-foot candles plus or minus some amount with localized lighting for more intense illumination at specific points. In simple terms, a good school lighting system lets students see visual tasks quickly, accurately, and comfortably. Since these tasks vary widely, a general system of illumination cannot make seeing equally fast and accurate at all times--or places. The state of Wisconsin requires a minimum of 30 plus or minus 10-foot candles of maintained illumination in the school classroom.

The amount of light is measured in foot-candles[1]. Until recent years, this unit of illumination was used almost exclusively as a measure of adequacy of school room lighting. Today it is recognized that this is only one factor in providing optimum light conditions:

Brightness (Foot-Lambert) = Intensity of Light (foot-candles) x Reflection Factor

Brightness: This is the product of intensity of illumination and the reflection factor. Brightness increases directly with increases in intensity of foot-candles. The task should be brighter than the background. Good contrast within the task is essential.

Intensity of Illumination: The need, here, is for variability with task needs. General illumination will be adequate if maintained at a reasonably low level, consistent with the principal tasks. Under a good system of glare-free illumination, for example, an experimental sample of handwriting with number two pencil on foolscap paper required 63 foot-candles of light for 99 percent visual accuracy. While the same handwriting sample using pen and ink required only 30 foot-candles of light for the same degree of performance. With printed materials, the size of the type, the gloss rating of the ink and the opaqueness and reflectance of the paper are factors in the visibility of the task and especially in the amount of light required to see it well.

---

[1] Foot-candle: This is a unit for measuring illumination. It is the amount of light produced by a standard candle at a distance of one-foot.

Reflection Factors: This is the ratio of the intensity of light falling on the surface to the light which is reflected. The amount of reflection will depend upon the color of the surface and the color of the light falling upon the surface. Reflectances of major surfaces in teaching-learning areas also affect visual environment. Dark colored surfaces absorb light. Lighter colored surfaces reflect and re-reflect light creating a better balance of brightness than is possible with darker surfaces.

Procedure

A research project funded by the U.S.O.E. (Cook and Blessing, 1970) had as part of its research activities the filming of the classrooms over a three-semester period. Light readings were obtained concomitant with the filming. The filming day was divided into four periods:

(1) Morning - before recess
(2) Morning - after recess
(3) Afternoon - before recess
(4) Afternoon - after recess

During the first semester, each of 23 classes were cycled through three of the periods. During the last two semesters, the 20 classes remaining in the project were cycled through all four periods.

ASA 160 T-stop readings were obtained from the light meter on an Eclaire camera. The T-stop readings were transformed to foot-candles.

The classes involved in the study were housed in schools located in or near the inner core of Milwaukee. With one exception the schools were at least 30 years old.

Results

In Table 1 are shown the average lighting levels of 23 EMR classrooms as a function of seasonal variations. Examination of the table reveals that seasonal variations do have a marked effect on the lighting levels in the classroom. During the late spring levels were around 84 foot-candles, while during the winter at the lowest point, they were 58 foot-candles. In general, it must be concluded that seasonal variations do have a marked effect on lighting in classrooms.

Table 1

Average lighting levels in 23 EMR classrooms as a function of seasonal variations

Time During Which Readings Were Taken	Lighting in Foot Candles
Feb. 2 to Feb. 14, 1967	70.
Mar. 20 to April 4, 1967	68.
May 15 to May 26, 1967	74.
Sept. 24 to Oct. 4, 1967	76.
Nov. 27 to Dec. 7, 1967	58.
Mar. 4 to Mar. 15, 1968	71.
May 20 to June 4, 1968	84.

In Table 2 is shown a frequency distribution of the average light levels of the 23 classes over the three semesters. Since some classes dropped out of the project, the averages shown are not all based on seven separate time periods. The state codes require maintained illumination within the range of 40-60 foot-candles. Using this criterion, only three of the classes had an average illumination level within this range. One fell below while the other 19 appeared to have a maintained illumination outside and above this range.

Table 2

Average Light Levels in 23 EMR Classrooms Over Three Semesters

Light Levels in Foot-Candles	Classroom Frequency
106	1
93	1
80-89	5
70-79	6
60-69	6
50-59	2
47	1
28	1

Discussion

This paper has not tried to be definitive, but rather provocative and to raise several questions. The most obvious question can address itself to the generality of the findings, not only to special education but also regular education classrooms. Do only 13% of the classes have a maintained illumination within the range of the state codes? More important, do 83% of the classrooms have illumination above the maximum required by the state codes? Even if the percentage is half this the findings do suggest that too much illumination whether due to too bright lights, too many windows, too many sources of glare and reflection, should be a concern of the individual teacher.

To what extent do lighting conditions affect the learning process in terms of motivation, sensory impact, etc? This question has hardly been tackled let alone answered. No doubt other environmental variables such as temperature, sound levels, crowding, etc., also influence an output such as learning rate. However, it should prove informative for educational decision making to determine the extent to which variations in illumination influence learning rate.

REFERENCES

Cawd., John J., Blessing, Kenneth R. "Class Size and Teacher Aides as Factors in the Achievement of the Educable Mentally Retarded" Final Report, Project No. 6-2620, 1970.

Gibson, Charles D. "Today's Concepts in School Lighting" *American School Board Journal*, June, 1965, 21-24.

Harmon, Darell B. *Lighting as Flexible as Your Classroom*, Vermilion: Wakefield Company, 1955, 40 pp.

Seagers, Paul W. "Lighting the Classroom," *NEA Journal*, 51, October, 1962, 30-31.

APPENDIX B

Test Manual and Scoring Materials for Behavior Analysis Scale - Educable (BASE)

BASE

EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE
ALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUC
HAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SC
E BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALY
LE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHA
SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE
YSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCA
OR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE
UCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR
LE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHA
AVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SC
SIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCAB

EXPERIMENTAL VERSION, JUNE 1969 EXPERIMENTAL VERSION, JUNE 1969 EXPERIMENTAL VERSION, JUNE 196

CABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR

RUCTION MR. WILLIAM C. KAHL, SUPERINTENDENT, WISCONSIN DEPARTMENT OF PUBLIC INSTRUCTION MR. WIL

OR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE

D CHILDREN MR. JOHN W. MELCHER, ADMINISTRATOR, DIVISION FOR HANDICAPPED CHILDREN MR. JOHN W. M

LE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEH
CABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR AN
ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – E
HAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SC
ABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANA
LE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHA
ALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDU
OR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE
UCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR A
E – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHA
ALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDU
A BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR AN
E – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHA
IS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCAB
OR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE
UCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR A
LYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDUCABLE BEHAVIOR ANALYSIS SCALE – EDU

BEHAVIOR ANALYSIS SCALE—EDUCABLE

Experimental Version, June 1969

by John J. Cook, Ph.D.

Coordinator, Research Design and Administration
Division for Handicapped Children
Wisconsin Department of Public Instruction

[illegible]

This instrument was developed as part of a project entitled, "Class Size and Teacher Aides as Factors in the Achievement of the Educable Mentally Retarded," funded by U.S.O.E. Grant Number OEG-3-6-062620-1879 Project Number 6-2620.

Principal investigators were:

- Kenneth R. Blessing, Ph.D., Director of the Bureau for Special Education, Wisconsin Department of Public Instruction
- E. Donald Blodgett, Ed.D., Executive Director of Exceptional Education, Milwaukee Public Schools
- Victor J. Contrucci, M.S., Program Chief - Mentally Handicapped Section, Wisconsin Department of Public Instruction
- John J. Cook, Ph.D., Coordinator, Research Design and Administration, Wisconsin Department of Public Instruction

The research reported herein was performed pursuant to a contract with the Office of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

Table of Contents

Page

Procedure 2

Results 3

Parameters of Scale Scores 7

Use of the BASE 8

List of Illustrations

Table 1 Means and Standard Deviations of Subjects' Age, I.Q. and WRAT Scores 2

Table 2 Item Loadings and Optimal Weights for the Five Steps of the BASE for Each of Five Factors or Sub-Scales 3

Table 3 Percentages of Total Variance and Common Variance Accounted for by the Five Factors 5

Table 4 Internal Consistency Reliability (r) and Standard Error (SE) of the BASE Sub-Scales When Administered on Two Occasions and in Terms of Composite Ratings (N=97) 6

Table 5 BASE Subtest Scores as a Function of Sex on Three Testing Occasions Using the Same Subjects 7

Table 6 Intercorrelations of the Scale Scores on the Two Testing Occasions 8

Figure 1 Parameters of Sub-Scales of BASE 9

Appendix 1 Behavior Rating Scale - Instructions and Questionnaire

Appendix 2 Transparency - Correction Sheet for Questionnaire

BEHAVIOR ANALYSIS SCALE-EDUCABLE

John J. Cook, Ph.D.

During the course of a project[1] funded by the U.S.O.E. and administered by the Wisconsin Bureau for Handicapped Children[2], a behavior rating scale was developed to serve as a dependent measure of behavioral changes in the classroom. The scale, developed for use with educable mentally retarded children, was assigned the name Behavior Analysis Scale-Educable or BASE. There is, however, no reason to believe the scale's use should be restricted to the mentally retarded. The content of the items suggest that it might be useful with any type of child above six or seven years of age who is attending school. The scale's function should probably be restricted to detecting gross changes over fairly extended periods of time or for the categorization of pupils at a given point in time.

The items for BASE were obtained in the following manner. William Gardner of the University of Wisconsin originally recognized the need for a behavior rating scale for the mentally retarded. He collected 130 statements descriptive of the behavior of the educable mentally retarded (EMR) from teachers of the EMR. Some 400 EMR students in the state were then rated by their teachers on these items. Item analysis subsequently reduced the number of items to 64. These 64 items constituted the scale as used in the project.

The purpose of the present paper is to present the results of analyses which explored the following areas:

1. The consistency of the factor structure of the instrument.

2. The optimal weighting of the five possible responses to each item as a function of the factor on which the item has the most unambiguous loading.

3. The reliability of the instrument when used by teachers and by teacher aides.

4. The validity of the instrument.

5. The derivation of a factor score for each subject using the weighting scheme.

6. The delineation of some parameters of the factor scores.

A copy of the test along with administration and scoring instructions accompanies this paper.

---

[1] Class Size and Teacher Aides as Factors in the Achievement of the Educable Mentally Retarded-Grant Number [illegible] Project Number [illegible]

[2] Division for Handicapped Children

Procedure

As part of the overall testing procedure in the aforementioned project, 23 EMR teachers and 12 teacher aides were asked to rate all the children in their classes on the BASE. In the latter part of January and early part of February, 1967, 375 children were rated. A child was not rated unless he had been known by the rater for at least a month. Two weeks after the initial rating, each teacher and each aide was asked to randomly select four children from their class and to rate them on the scale again. Response was not complete, but 86 children were re-rated by teachers and 44 by aides. These re-ratings provided the data for the reliability analyses.

In November of 1967, the BASE was again administered to all children in the project. Of the 352 children rated, 239 were in the original group rated. Attrition was due to transfers to secondary school, out-migration from the area, death, and, most important, replacement teachers not wanting to get into the project.

Table 1

Means and Standard Deviations of Subjects'
Age, I.Q. and WRAT Scores

	February 1967			November 1967		
	N	$\bar{X}$	SD	N	$\bar{X}$	SD
C.A.	375	120.08	22.34	352	116.86	21.49
I.Q.	374	67.37	7.44	352	68.21	7.10
Reading	365	33.79	13.29	342	34.51	13.35
Spelling	365	24.23	7.26	342	24.08	6.89
Arithmetic	365	20.52	6.48	342	21.42	5.66

Table 1 shows the salient characteristics of the two groups.

The subjects comprising the first and second ratings were very similar. Means and standard deviation are quite similar. The reduction in C.A. of three months was probably due to the older children moving out of the immediate classes into high school classes. Differences in N were due to an inability to test all the children on all the instruments. A final rating of all pupils, done in May of 1968, does not enter into the analysis reported herein except for the determination of the final weights.

Both sets of teacher rating data and Gardner's original data were factor analyzed. The factoring procedure used was an image analysis followed by a varimax rotation of the resulting factors.[3] Comparative item loadings on the five factors were remarkably similar for the analyses lending support to the validity of the factor structure. Each item which had a substantial loading (greater than .40) was then assigned to one of five scales, corresponding to the five factors, on the basis of its largest loading.

Responses to each of the items in each of the five scales were then reweighted to maximize the internal consistency reliability of each scale. The exact procedure used was the method of reciprocal averages which has been programmed for the University of Wisconsin Computing Center. The method assumes that a single variable underlies all items in the scale. A priori assigned weights initiates an iterative procedure that converges to a weighting scheme which optimizes the internal consistency of the scale. The obtained set of weights were used to calculate scores for each person on each of the five scales.

---

[3] This is a standard option of the University of Wisconsin Computing Center's factor analysis program

Separate RAVE analyses done on both the first and second administrations of the BASE produced item weights that were virtually identical. To conserve on time and effort, a composite RAVE analysis was done using all student ratings from all three project testing sessions. Thus, the computer derived the weighted score for each individual on each of the five scales for each of the three occasions. The weights from this composite analysis were also virtually identical to those derived from the first two analyses.

Results

Tables 2 and 3 present the results of the image factor analysis and the RAVE-derived weights for each item on the corresponding scales. (Scales rather than factors is the preferred term and will be used hereafter.) The empirically derived labels for each of the five scales are:

Scale 1. Two features seem to underlie the behavior described by the 26 items loading on this scale. Aggressiveness is one, with the other being the negative denotation of the behavior described. This scale was labeled **Disruptive Behavior.**

Table 2

Item Loadings and Optimal Weights for the Five Steps of the BASE for
Each of Five Factors or Sub-Scales

Factor I: Disruptive Behavior

Item		Loading	Weight
29.	gives other pupils dirty looks	-.84	12345
8.	threatens to hurt other people when angry	-.83	12345
10.	is stubborn	-.83	12345
57.	has difficulty in controlling temper—is easily upset	-.83	12345
30.	has to have everything his own way	-.81	12345
54.	when criticized by teacher frequently reacts in sullen manner—sulks or pouts	-.80	12345
33.	uses abusive language toward others	-.79	12355
56.	has a "chip on his shoulder"	-.79	12355
11.	companions find it difficult to get along with him	-.77	12345
3.	if passed over or slighted or if things go wrong, is inclined to sulk	-.76	12345
20.	is uncooperative	-.74	12355
50.	does not respect other people's belongings	-.74	12345
2.	is disruptive and boisterous	-.73	12345
14.	uses profanity	-.73	13345
22.	when angry will refuse to speak to anyone	-.71	12345
48.	finds fault with instructions given by adults	-.71	12355
17.	resents even the most gentle criticism of his work	-.70	12345
46.	is discontented with everything in general	-.68	12455
49.	acts up when I'm not watching	-.68	12234
16.	has changeable moods	-.67	12345
42.	does not forget things that anger him	-.67	12344
41.	seems unconcerned when he misbehaves	-.64	12345
27.	shows a dislike for school	-.57	12445
40.	"drags his feet" when requested to do something	-.55	12344
64.	is upset by criticism or disapproval	-.51	12334
63.	is considerate of others	.44	44321

Factor II: Attentive Involvement in Classroom Activities

1.	is able to concentrate on things	.82	12345
45.	can be trusted to carry out instructions promptly and attentively	.81	12345
15.	finishes his classroom assignments	.80	12345
31.	works well by himself	.80	12345
4.	is interested in schoolwork	.79	12345
5.	is alert in class	.78	12345
24.	can be trusted with responsibility	.75	12345
52.	is a helpful person	.70	12345
55.	readily becomes enthusiastic over a new project or activity	.56	12345
62.	volunteers to help when there is work to be done	.56	12345
34.	gives up when faced with difficulty	-.52	54321
63.	is considerate of others	.51	12345
8.	is popular with classmates	.47	12345

Factor III: Timidity

13.	is timid in meeting people	.79	12345
38.	is afraid of strange adults	.76	12355
12.	becomes embarrassed easily	.74	12345
35.	feels shy when the center of attention	.73	12345
44.	becomes frightened easily	.68	12355
7.	displays feelings of inferiority	.47	12344
36.	has a hard time making decisions	.41	12344

Factor IV: Succorance

60.	seeks adult attention	-.76	11345
39.	seeks affection from others	-.72	12345
51.	seeks out adult help in doing things	-.71	12345
58.	seeks approval for his behavior	-.61	12345
6.	demands the attention of the teacher and pupils	-.51	12345
32.	tries to arouse the sympathy of others	-.49	22355
59.	acts dependent	-.47	12344

Factor V: Withdrawal

23.	seems to be drowsy	.80	12355
21.	is tired and listless—little energy	.79	12345
61.	is inactive or sluggish	.75	12345
10.	maintains same facial expression (blank stare or far-away look)	.69	12345
28.	daydreams	.68	12345
26.	appears to be preoccupied	.67	12345
18.	seems to be off in a world of his own	.61	12345
25.	displays little interest in his environment	.61	12355
47.	prefers to be alone—works and plays alone	.57	12344
53.	appears to be sad, unhappy, lonely	.56	12344
43.	does not participate in group activities—stays in background	.49	12345
37.	shows little interest in classroom routine	.46	12345

Scale II. The two notions of involvement and attentiveness relative to the classroom situation characterize the items loading on this scale. Thus, the label **Attentive Involvement in Classroom Activities** seems appropriate.

Note that there are both positive and negative loadings of items within the scale. Item 34, "gives up when faced with difficulty" has a negative loading while all the others have positive loadings. As can be seen from the weights presented, the RAVE program reverses these for consistency in interpretation. Thus, category 1 "never, does not apply" is weighted 5 on item 34, whereas on all other items on this scale it is weighted 1.

Scale III. Shyness with elements of withdrawal seems to characterize the items on this scale. The label which suggests itself is **Timidity.**

Scale IV. This scale was called **Succorance** where succorance is defined as "the tendency to solve one's problems by seeking aid or protection from someone else."

Scale V. This scale seems to be the polar opposite of Factor II, **Attentive Involvement.** Some schizoid elements seem present, but in order to make the label descriptive yet value neutral, it is proposed that this scale be called that of **Withdrawal.**

Table 3

Percentages of Total Variance and Common Variance
Accounted for by the Five Factors

	% of Total Variance	% of Common Variance
I. Disruptive Behavior	22.04	33.69
II. Attentive Involvement in Classroom Activities	12.81	19.27
✓ III. Withdrawal	6.14	9.39
III. IV. Timidity	5.03	7.69
V. Succorance	10.30	16.04
TOTAL	56.32	86.08

The percentages of total and common variance accounted for by each of the factors is noted in Table 3. These percentages vary from 22.04 to 5.03 of the total variance and from 33.69 to 7.69 of the common variance. Together the five factors account for 56.32% of the total variance and 86.08% of the common variance. As the decision rule was made to include only those items unambiguously loading on a factor at .40 or greater, the sum of squares of loadings of items in each factor do not sum to indicated percentages of total and common variance. Of the 64 items entering into the analysis, all are represented in the five factors, with one item, 63, contributing to both Scales I and II.

Table 4 presents the internal consistency reliability estimates for each of the five optimally weighted scales. As should be expected, these reliabilities vary with the number of items in the scale, with Scales I, II and V all having r's between .92 and .97. Scales III and IV, the two seven-item scales, have reliabilities between .80 and .86.

That these figures tend to be high is demonstrated by comparisons with the results of the rating-rerating data. Eighty-six children were re-rated within a three-week period by teachers with an r of .601, while 44 re-ratings by aides produced an r of .515 with the initial rating.

Table 4

Internal Consistency Reliability (r) and Standard Error (SE) of the BASE Sub-Scales When Administered on Two Occasions and in Terms of Composite Ratings (N = 87)

Scales		Feb. 1967		Nov. 1968		Composite		No. of
		r	SE	r	SE	r	SE	Items
I.	Disruptive Behavior	.96	4.47	.97	4.20	.97	3.80	26
II.	Attentive Involvement in Classroom Activities	.94	3.42	.95	3.09	.93	2.44	13
III.	Timidity	.85	2.25	.86	2.14	.85	1.70	7
IV.	Succorance	.80	2.01	.83	1.84	.82	1.83	7
V.	Withdrawal	.92	2.90	.93	2.67	.92	2.40	12

Basically, the question of validity revolves around the question of whether or not one is measuring what one thinks he is measuring. Three types of validity have been noted-content, predictive or concurrent and construct validity. A consideration of the validity of the BASE in terms of these types follows.

The content validity of the BASE is ensured by the manner in which the subtests were derived, i.e., Gardner's original selection.

A form on construct validity can be demonstrated if we can substantiate certain hypotheses regarding expected differences in subtest scores. That is, we would hypothesize that boys score higher than girls on Scale I, *disruptive behavior*, and that girls score lower than boys on Scale III, timidity. Sex differences of this nature have been regularly demonstrated with other behavior rating scales.

These sex differences were investigated for the BASE. In addition, the changes in subtest scores were explored over the course of three semesters. Only those subjects who were given the BASE on all three occasions were used. A total of 109 males and 69 females are represented on the three occasions. Salient characteristics of the two groups at the beginning of the first semester were:

		CA(mo.)		IQ	
	N	$\bar{X}$	SD	$\bar{X}$	SD
Male	109	117.4	20.9	68.5	6.6
Female	69	113.7	19.6	66.8	6.9

The CA and IQ differences were not statistically significant at the .05 level (F = 1.36 and 3.37 respectively with df = 1 and 176).

Table 5 (a,b,c,d,e) shows the means and standard deviations of the five subtests as a function of sex over time. Significant differences between the means are indicated by an asterisk and are based on a step-wise discriminant analysis. Girls were rated as significantly more timid than boys on all three occasions. Boys manifested more disruptive behavior than the girls on all three occasions, but statistical significance was obtained on the first and second testings only. Attentive involvement in classroom activities favored the girls throughout but was statistically significant on the final testing only.

Changes over time were minimal for Scales III, IV, and V. Scale II, attentive involvement, showed slight increases over time for both sexes. Scale I, disruptive behavior, demonstrated marked increases for both sexes over time. Most of this increase occurred from the first to the second testing, i.e., March, 1967, to September, 1967. Since most of the classrooms were located in the inner core of Milwaukee, it is not unlikely that the racial troubles of that summer were manifest in the behavior of the children in the classroom.

Evidence for predictive or concurrent validity is not available at the present time.

Table 5

BASE Subtest Scores as a Function of Sex on Three Testing
Occasions Using the Same Subjects

a) Scale I. Disruptive Behavior

		T_1	T_2	T_3
Male	$\bar{X}$	60.5	66.9	66.9
	SD	20.3*	18.1*	20.3
Female	$\bar{X}$	58.7	60.7	63.3
	SD	19.0	15.2	17.3

b) Scale II. Attentive Involvement

		T_1	T_2	T_3
Male	$\bar{X}$	38.5	40.3	41.2
	SD	9.5	8.5	9.3*
Female	$\bar{X}$	40.6	41.9	44.0
	SD	9.9	10.0	9.2

c) Scale III. Timidity

		T_1	T_2	T_3
Male	$\bar{X}$	17.8	17.4	18.3
	SD	5.2*	4.2*	4.8*
Female	$\bar{X}$	20.1	20.6	20.9
	SD	5.4	5.2	5.4

d) Scale IV. Succorance

		T_1	T_2	T_3
Male	$\bar{X}$	19.8	19.8	20.0
	SD	5.0	4.6	4.9
Female	$\bar{X}$	19.8	20.5	20.2
	SD	4.8	4.4	4.1

e) Scale V. Withdrawal

		T_1	T_2	T_3
Male	$\bar{X}$	28.5	28.8	30.5
	SD	8.9	7.9	7.9
Female	$\bar{X}$	29.3	29.7	30.2
	SD	9.7	9.7	8.8

*Significant at .05 Level

Parameters of Scale Scores

In this section, an attempt is made to establish some boundaries between which scores on the subtests of the BASE can be expected to fall with a known probability. That is, an attempt is made to provide baseline data from which users of the scale can compare their own ratings.

During the course of the project some 965 behavior rating scales were administered. These scales were administered by one teacher to one child on three occasions, by one teacher to different children on three occasions and by different teachers to different children on different occasions. In spite of this heterogeneity of administration and subjects, the means and standard deviations of the five subtests do not deviate markedly from the data shown in Table 5 when the same children were rated by the same teachers.

Figure 1 presents means and upper and lower bounds of the test score based on a plus or minus one standard deviation spread. This information is not meant to be definitive since the population on which it is based is not representative of all EMR's in either the state or the county. However, since the instrument was a research instrument and little previous information is available, the data presented can be considered a point of departure in norm development should the instrument prove of value in the field.

Table 6 presents the intercorrelations among the factor scores on the five factors. The intercorrelations of the factors on the two occasions remained substantially the same. In general, the four scales (I, III, IV and V) with a negative connotation tended to intercorrelate high and positive and to correlate negatively with the one scale (II) having a positive connotation. On both occasions, the consistently independent scales were Disruptive Behavior (I) and Timidity (III).

Table 6

Intercorrelations of the Scale Scores on
the Two Testing Occasions

a) Feb., 1967

Scale	I	II	III	IV	V
I. Disruptive Behavior	—	-.61*	.09	.57*	.33*
II. Attentive Involvement		—	-.36*	-.14*	-.74*
III. Timidity			—	.12	.60*
IV. Succorance				—	.19*
V. Withdrawal					—

b) Nov., 1967

Scale	I	II	III	IV	V
I. Disruptive Behavior	—	-.67*	.07	.58*	.38*
II. Attentive Involvement		—	-.38*	-.21*	-.75*
III. Timidity			—	.23*	.62*
IV. Succorance				—	.31*
V. Withdrawal					—

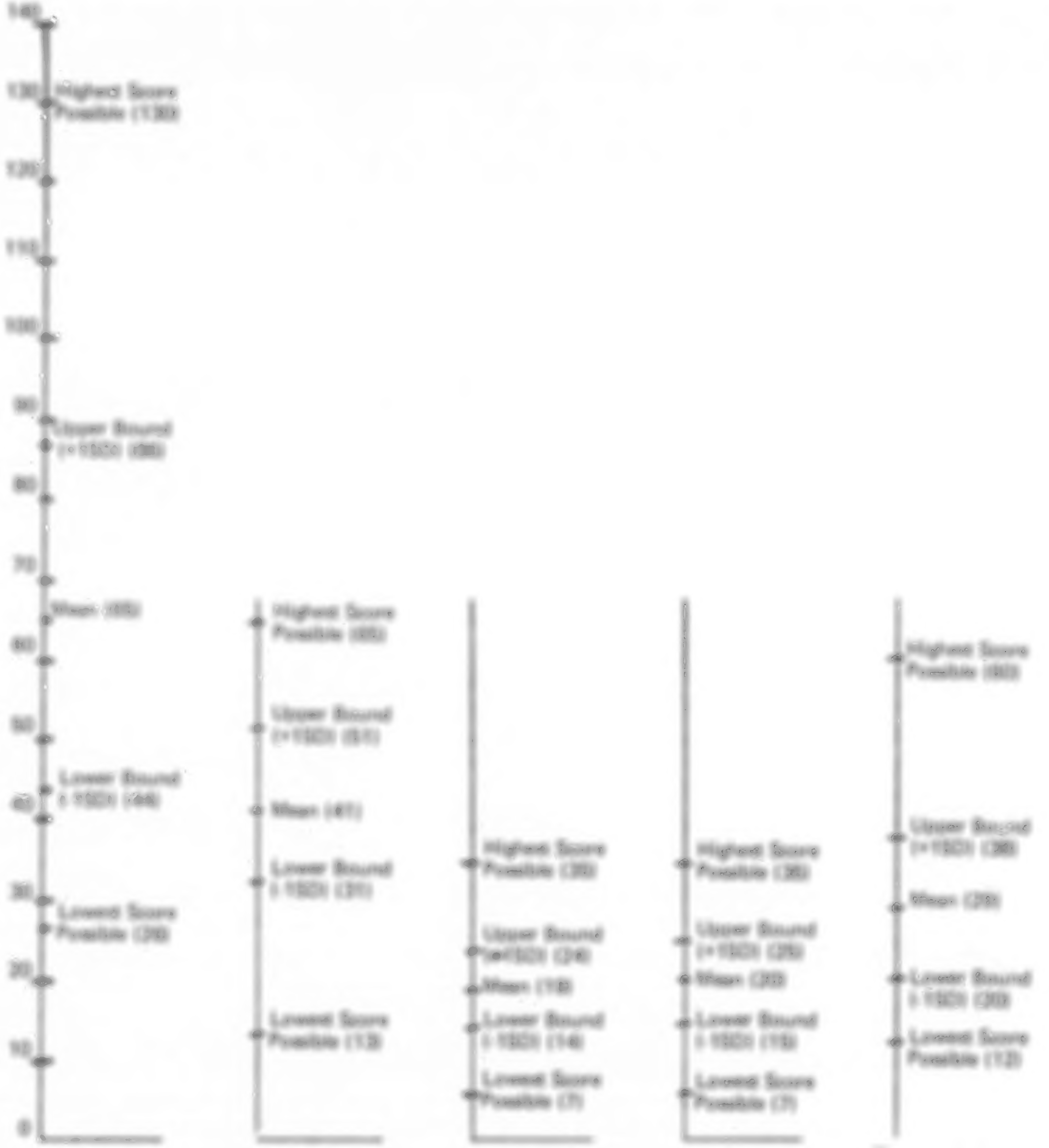
*Significant at the .01 level of confidence

Use of the BASE

It would seem that the BASE would have merit as an evaluative instrument insofar as the individual teacher or broader program has as its goals the modification of behaviors defined by the items comprising the five subtests or factors of the scale. These item clusters or factors have been labeled as **Disruptive Behavior, Attentive Involvement in Classroom Activities, Timidity, Succorance** and **Withdrawal**. The instrument might also, in conjunction with other information, be used as a screening device for placement purposes. It could also be used as a means of testing descriptive hypotheses with respect to the various subtests.

A copy of the BASE as used in the present project accompanies this paper. The instructions are included for its use. It is suggested that the scale, if used in its full form, be used as shown for the subscale items are sufficiently dispersed to prevent any bias that might occur if they were grouped together. After administration of the scale, the item weights for each subtest shown in Table 3 should be used to obtain total scores for each of the subtests. The quasi-standardization data in Figures 1-5 can be used for comparing one child or a group of children with our larger samples of EMR's having characteristics similar to those depicted in Tables 1 and 2. In general, the hypothesis might be entertained that if the child's score exceeds the upper, or lower bound, depending upon the subtest, then this aspect of his behavior might be given further consideration.

Profiles based on the five subtests can be generated by using standard scores. Once again the child under consideration should be comparable to the norming group.



Scale 1
Parameters of the Disruptive Behavior Scale of BASE.

Scale 2
Parameters of the Attentive Involvement Scale of BASE.

Scale 3
Parameters of the Timidity Scale of BASE.

Scale 4
Parameters of the Succorance Scale of BASE.

Scale 5
Parameters of the Withdrawal Scale of BASE.

Figure 1 Parameters of Sub-Scales of BASE

BEHAVIOR RATING SCALE - EDUCABLE B.A.S.E.

NAME ______________________ DATE ______________________ SUBJECT CODE# _____

SEX ______________________ TEACHER ______________________ CLASS __________

BIRTHDATE ______________________ SCHOOL ______________________ CELL NUMBER# _____

RATING GUIDE

1. Base rating on pupil's behavior since the beginning (end) of the semester.

2. Base your rating on your own experience with the pupil being rated.

3. Complete each item without consideration of ratings on other items.

4. Base your ratings on observations of outward behavior and not on what you may think may underlie this behavior.

5. Do not spend more than a few seconds on any item. Rate the item quickly.

6. Complete all items.

7. Each item is to be rated with a 1, 2, 3, 4, or 5 according to the following:

   1 = Never describes
   2 = Rarely describes
   3 = Occasionally describes
   4 = Often describes
   5 = Very frequently describes

SCORING INSTRUCTIONS

Scoring keys are provided to facilitate scoring of the tests. The steps in scoring are:

1. Place the transparent scoring key on the test page. Appropriate positioning of the key is obtained by lining up the question numbers on the key with the numbers on the test.

2. Each question must be scored on one of five scales identified by these designs– Scale I ▲▲ , Scale II ●● , Scale III ●● , Scale IV ●● and Scale V ●● .

3. At the end of each question on the key are five boxes corresponding to the five responses on which each item was checked. In the upper left corner of each box is a series of weights ranging from one through five.

4. With the scoring key in position, add up the weights checked by the rater of all questions marked ▲▲ . Enter this sum in the box under Scale I on the bottom of this page. Then add up the total weights of all questions marked ●● and enter the sum under Scale II.

   Please note that Question No. 63 applies to Scale I and Scale III. The weights for Scale I are in the upper left corner of the five response boxes. The weights for Scale III are in the upper right corner of the same five response boxes.

5. Continue this process for the remaining three scales entering the sum each time in the appropriate box on this page.

TOTAL TEST SCORE

☐	☐	☐	☐	☐
▲ Scale I	● Scale II	● Scale III	● Scale IV	● Scale V

Behavior Rating Scale —
Educable (B.A.S.E.)
PL-(04C)-SE-70

	Never describes 1	Rarely describes 2	Occasionally describes 3	Often describes 4	Very frequently describes 5
1. is able to concentrate on things					
2. is disruptive and boisterous					
3. if passed over or slighted or if things go wrong, is inclined to sulk					
4. is interested in schoolwork					
5. is alert in class					
6. demands the attention of the teacher and pupils					
7. displays feelings of inferiority					
8. is popular with classmates					
9. threatens to hurt other people when angry					
10. is stubborn					
11. companions find it difficult to get along with him					
12. becomes embarrassed easily					
13. is timid in meeting people					
14. uses profanity					
15. finishes his classroom assignments					
16. has changeable moods					
17. resents even the most gentle criticism of his work					
18. seems to be off in a world of his own					
19. maintains same facial expression (blank stare or far away look)					
20. is uncooperative					
21. is tired and listless-little energy					
22. when angry will refuse to speak to anyone					

	Never describes 1	Rarely describes 2	Occasionally describes 3	Often describes 4	Very frequently describes 5
23. seems to be drowsy					
24. can be trusted with responsibility					
25. displays little interest in his environment					
26. appears to be preoccupied					
27. shows a dislike for school					
28. daydreams					
29. gives other pupils dirty looks					
30. has to have everything his own way					
31. works well by himself					
32. tries to arouse the sympathy of others					
33. uses abusive language toward others					
34. gives up when faced with difficulty					
35. feels shy when the center of attention					
36. has a hard time making decisions					
37. shows little interest in classroom routine					
38. is afraid of strange adults					
39. seeks affection from others					
40. "drags his feet" when requested to do something					
41. seems unconcerned when he misbehaves					
42. does not forget things that anger him					
43. does not participate in group activities, stays in background					

	Never describes 1	Rarely describes 2	Occasionally describes 3	Often describes 4	Very frequently describes 5
44. becomes frightened easily					
45. can be trusted to carry out instructions promptly & attentively					
46. is discontented with everything in general					
47. prefers to be alone - works and plays alone					
48. finds fault with instructions given by adults					
49. acts up when I'm not watching					
50. does not respect other people's belongings					
51. seeks out adult help in doing things					
52. is a helpful person					
53. appears to be sad, unhappy, lonely					
54. when criticized by a teacher he quietly reacts in a sullen manner					
55. readily becomes enthusiastic over a new project or activity					
56. has a "chip on his shoulder"					
57. has difficulty in controlling temper - is easily upset					
58. seeks approval for his behavior					
59. acts dependent					
60. seeks adult attention					
61. is inactive or sluggish					
62. volunteers to help when there is work to be done					
63. is considerate of others					
64. is upset by criticism or disapproval					

Behavior Rating Scale -
Education (B.A.S.E.)
[illegible]

1. .... 1 2 3 4 5
2. .... 1 2 3 4 5
3. .... 1 2 3 4 5
4. .... 1 2 3 4 5
5. .... 1 2 3 4 5
6. .... 1 2 3 4 5
7. .... 1 2 3 4 5
8. .... 1 2 3 4 5
9. .... 1 2 3 4 5
10. .... 1 2 3 4 5
11. .... 1 2 3 4 5
12. .... 1 2 3 4 5
13. .... 1 2 3 4 5
14. .... 1 2 3 4 5
15. .... 1 2 3 4 5
16. .... 1 2 3 4 5
17. .... 1 2 3 4 5
18. .... 1 2 3 4 5
19. .... 1 2 3 4 5
20. .... 1 2 3 4 5
21. .... 1 2 3 4 5
22. .... 1 2 3 4 5

23. ...... 1 2 3 4 5
24. ...... 1 2 3 4 5
25. ...... 1 2 3 4 5
26. ...... 1 2 3 4 5
27. ...... 1 2 3 4 5
28. ...... 1 2 3 4 5
29. ...... 1 2 3 4 5
30. ...... 1 2 3 4 5
31. ...... 1 2 3 4 5
32. ...... 1 2 3 4 5
33. ...... 1 2 3 4 5
34. ...... 5 4 3 2 1
35. ...... 1 2 3 4 5
36. ...... 1 2 3 4 5
37. ...... 1 2 3 4 5
38. ...... 1 2 3 4 5
39. ...... 1 2 3 4 5
40. ...... 1 2 3 4 5
41. ...... 1 2 3 4 5
42. ...... 1 2 3 4 5
43. ...... 1 2 3 4 5

44. ...... 1 2 3 4 5
45. ...... 1 2 3 4 5
46. ...... 1 2 3 4 5
47. ...... 1 2 3 4 5
48. ...... 1 2 3 4 5
49. ...... 1 2 3 4 5
50. ...... 1 2 3 4 5
51. ...... 1 2 3 4 5
52. ...... 1 2 3 4 5
53. ...... 1 2 3 4 5
54. ...... 1 2 3 4 5
55. ...... 1 2 3 4 5
56. ...... 1 2 3 4 5
57. ...... 1 2 3 4 5
58. ...... 1 2 3 4 5
59. ...... 1 2 3 4 5
60. ...... 1 2 3 4 5
61. ...... 1 2 3 4 5
62. ...... 1 2 3 4 5
63. .... [illegible]
64. ...... 1 2 3 4 5

APPENDIX C

Language Analysis

1. Analysis of Language Samples
2. Factors Affecting Language Indices
3. Changes in Language Indices as a Function of Race, Sex, and IQ

1. ANALYSIS OF LANGUAGE SAMPLES

John J. Cook, Ph.D.
Carol Tuck
Carol Sexton, Ph.D.

The assumption has been made in this project that if one selects a group of teachers at random, their collective activity during the school year has a certain commonality with respect to the direction of their activities which, for lack of a better term, we now call their "instructive endeavors." These instructive endeavors have an effect upon the extent to which the child partakes of, or becomes involved in, education, in turn defined as "progressive or desirable changes in a person as a result of teaching and study." (English and English, 1958). These "progressive and desirable changes" can be considered to occur in many areas of functioning, particularly when one considers special education of the mentally retarded, physically handicapped, deaf, blind, emotionally disturbed, and those children with special learning disabilities.

Many factors--environmental, administrative, economic, and sociopsychological---can influence the instructive endeavors of the group of teachers. In the present study, two factors were manipulated--class size and use or non-use of a paraprofessional in the classroom--with primary (C.A. 7-10) and intermediate (C.A. 10-13) children in special classes for the educable mentally retarded (EMR). It is also possible to define a number of "progressive and desirable changes" which should accrue as the result of the instructive endeavors of the group of teachers. We defined three broad areas of change:

1) Academic--indexed by the WRAT with its subtests of reading, spelling, and arithmetic.

2) Behavioral--indexed by a teacher-administered behavior rating scale with five subtests determined by factor analysis.

3) Language--indexed by the Illinois Test of Psycholinguistic Abilities (ITPA) and the analysis of samples of spoken language.

The present paper is a description of how the language samples were obtained, transcribed, and analyzed. In obtaining a language sample, a situation must be set up so that an adequate sample of the child's language-encoding capability can be recorded. The stimulus for such a situation should be of high interest value, relatively uncomplicated--so that the listing of objects can be avoided--and fairly unstructured so that the stimuli can be talked about from a number of points of view. Non-captioned Charles Addams drawings appeared to meet these criteria. Subsequent field testing with eight of the drawings, using EMR children, revealed that approximately three minutes of storytelling about and description of the drawing was often obtainable with the presentation of one picture, and very nearly always so with the successive presentation of two drawings.

Accordingly, eight uncaptioned Charles Addams drawings were reproduced with permission. Field testing allowed a preference rating of the eight stories. The four most preferred stories were then chosen and used in various combinations on the three testing times.

Obtaining Speech Samples

An example of the instructions given to the testers picking up the speech samples is attached in Appendix B. A fairly rigid, standardized approach to obtaining the speech samples was attempted initially. However, it soon became apparent that some difficulty was encountered in getting certain children to respond fully. By loosening up the procedure and allowing more interplay between the tester and the child, this difficulty was in large measure obviated. The speech samples were recorded monaurally, in most instances at 3 3/4 ips.

Transcription of Speech Samples

Speech samples were obtained from some 170 [illegible] children. About 40% of these children were Negroes from the inner-city area of Milwaukee. Thus, dialectal problems loomed large in transcription.

A. Linguistic Considerations

The speech samples were transcribed from the tapes both in ordinary English orthography and in a semi-phonemic transcription, a phonemic transcription based on Trager and Smith, (1951) but adapted to fit the symbols present on a standard English typewriter. All false starts, unfinished words, repetitions, hesitations and "fillers" like *uh* and *um* were included in the transcript. Every effort was made to have the typed transcript represent the subjects' speech faithfully. These hesitation phenomena were not, however, subdivided as by Maclay and Osgood (1959) or Levin and Silverman (1965).

1. Pauses and intonation contours were carefully indicated. Pauses and breaks in the speech were marked with slashes: a short break or hesitation was indicated by a single slash (/), and a longer pause by a double slash (//). Extremely long silences were duly noted by the transcriber, who typed the comment "long pause" between utterances separated by long periods of silence.

2. Intonation patterns (or the pitch melodies of utterances) are closely tied to the occurrence of major sentence-stresses in English in that major stresses are generally accompanied by relatively high pitch. Pre-pause pitch movements, however, are quite independent of stress, and were indicated in the transcription by the following punctuation marks:

   a marked rise in pitch was represented by a question mark (?)

   a slight rise in pitch, by a colon (:)

   a marked fall in pitch, by a period (.)

   a slight fall in pitch, by a comma (,)

Another aspect of intonation--tempo--was indicated when the last word or phrase before a pause was drawn out or lengthened. In this study a dash (--) represented this slowing up in the tempo of speech. It was used in conjunction with the punctuation marks listed above when the pitch changed; when a steady pitch was maintained while the final word was drawn out, only the dash was used.

Examples: this man--./ he look like he put in/ some cement./ an' it's a shovel--/ it's a long stick--.//

3. English consonants were represented by the following symbols, listed with word-initial, medial and final examples (where applicable):

			phonetically
Stops			
voiceless	p	pit, supper, tip	pit, səpər, tip
	t	tick, litter, hit	tik, litər, hit
	tsh	chip, pitcher, witch	tship, pitshər, witsh
	k	cap, fickle, neck	kæp, fikəl, nek
voiced	b	bit, rubber, rib	bit, rəbər, rib
	d	dip, rudder, grid	dip, rədər, grid
	j	jump, fragile, ridge	jəmp, fræjəl, rij
	g	gun, regal, dig	gən, riygəl, dig
Fricatives			
voiceless	f	fat, muffle, rough	fæt, məfəl, rəf
	th	thick, breath	thik, breth
	s	sit, assent, brace	sit, əsent, breys
	sh	shot, nation, cash	shat, neyshən, kæsh
voiced	v	vat, over, stove	væt, owvər, stowv
	dh	that, other, breathe	dhæt, ədhər, briydh
	z	zip, frizzle, phase	zip, frizəl, feyz
	zh	azure, rouge, measure	æzhər, ruwzh, mezhər
Nasals	m	mat, summer, room	mæt, səmər, ruwm
	n	nap, done, run	næp, dən, rən
	ng	singer, thing, strong	singər, thing, strong
Liquids	l	let, allow, all	let, əlaw, ɔl
	r	run, airy, car[1]	rən, eriy, ar

[1] For "syllabic" l, r, n, as in bottle, butter, button, see 5 below

When a nasal consonant occurred before k or g, it was written n, although pronounced like the ng of thing. Thus think and thing were transcribed as think and thing, and finger and singer were transcribed as finger and singer.

semivowels	w	with, swill	with, swil
	y	you, use, few	yuw, fyuw

4. Fourteen English vowel sounds were represented—eight simple vowels, three diphthongalized vowels, and three diphthongs. They are listed below with word-medial, initial and final examples (where applicable):

Simple	a	hat, apple	Phonetically:	[illegible]
	ah	pot, father, ma		[illegible]
	aw	bought, awful, raw		[illegible]
	ə	but, upper, America		[illegible]
	e	bet, ever		[illegible]
	i	bit, infant		[illegible]
	o	boat, over, row		[illegible]
	u	put, bush, book		[illegible]
Diphthongalized				
	ey	bait, ache, day		[illegible]
	iy	beat, eel, fee		[illegible]
	uw	boot, ooze, two		[illegible]
Diphthongs				
	ay	bite, ice, rye		[illegible]
	aw	bout, owl, now		[illegible]
	oy	boil, oyster, boy		[illegible]

5. "Syllabic" l, r, n were written as ul, ər, in when the vowel coloring of the syllables corresponded to the vowels u, ə, and i respectively. For some dialects, at least, the following examples are valid:

ul	bottles, ruffled, rattling	[illegible]
ər	butter, farmers, bird	[illegible]
in	buttons, mountain, softened	[illegible]

6. Stresses were also marked in the transcripts. Three or four degrees of stress have been established for English, but for this study it was considered sufficient to mark two degrees of major sentence stress by underlining and to leave weaker stresses unmarked. Generally there is one primary stress between two syntactically appropriate pauses, but the number of secondary stresses varies considerably from speaker to speaker.

By convention, if the primary stress occurred as the last stress before a pause, both it and all previous (secondary) stresses were underlined with a single line:

an' this man has shovels./

an dis man has shovels./

But if the primary stress was followed by one or more secondary stresses before the pause, it was underlined with a double line and all the previous and subsequent secondary stresses were underlined just once:

an' there's some trees in the background.//

an dəz sam triyz in ə [illegible]//

Some fifty orth-phonemic transcriptions were recorded, but the procedure was then abandoned because of time considerations.

3. Process of Transcribing

The first time through the transcriber typed the speech sample in ordinary English orthography, including in single parentheses the examiner's remarks and interjections and in double parentheses the transcriber's comments, e.g., ((long silence)) ((giggle)) ((out of breath)). When present in the speech sample, contractions were noted so that the transcriber could match the precise sample, e.g., wanna, gonna, gotta, lotta, won't, he's, there's. Incomplete words were included and written with a hyphen following to indicate that they were not complete. For example, even if the entire words dog and they were identifiable in context, only that portion of the word actually spoken was recorded: d- the-. Intonation marks, especially before pauses, were also recorded, as were the pauses. The speech samples were then timed to the nearest five seconds, and the time converted to decimals.

As the transcriber listened to the tape again, she marked from the orth-phonemic representation of what she had heard. Special attention was paid to missing sounds (t of in, like; d of and, around) and to substituted sounds (d for th in there's; n for th in the following; [illegible] at the end of the preceding word, etc.). Tense changes, intonation markers and pauses were then added from the transcription in English orthography and secondary stresses were indicated by underlining. Six spaces of tape were left between lines of type the first time through. The second time through the phonological representation was inserted two lines below the English orthography.

The analysts found the following procedure helpful. They would listen to the tape as far as a pause and then stop, trying to recall the first segment word for word. They could then guess at dubious items, but check these by listening to the tape a second time before typing. This listen-guess-check-type process was then repeated for the next phrase.

Analysis of Transcripts

A total of 37 tapes of two [illegible] were obtained from the transcripts. From these measures were derived indices which research and a related preventive [illegible] indicated were related to, or indices of, "reinforcing and [illegible] changes" in language-acquisition. Following are 1) operational definitions of the two terms, and 2) a tentative interpretation of the derived indices:

4. Definition of Terms

Time

Time was defined as the amount of total time of the utterances in the transcript minus the amount of time of the utterances of the tester. What remains to be counted as "time," then, is the period of time during which the subject himself was speaking. Any portion of a silence which ran over 15 seconds was subtracted from the time estimate. The time was expressed in minutes and proportions of a minute in units of five seconds, using the following table.

Seconds	Proportion of Minute
0	0.00
5	.08
10	.17
20	.33
25	.42
30	.50
35	.58
40	.67
45	.75
50	.83
55	.92
60	1.00

Words

The "word" was not used as a basic unit in the analysis because of the widely recognized problems involved in arriving at an operational definition. See the unit "token" below for an approximation of the concept of "word."

Syllables

These are the smallest unit of spoken language considered in the analysis. Syllables were counted primarily as a rough temporal index of the relative amount of time spent in mazes (a term defined below).

Tokens

A token was defined as any meaningful unit whose meaning is made clear by its context, if not by its phonological pattern, and which is separated from other likely tokens by silence. This means that contractions, such as can't and there's, and slurred word groups such as could've as in "I could've done that" were counted as one token each. By convention, phonological units within mazes--meaningful or not--were not counted as tokens.

The actual form of the transcription on the page will correlate directly with the boundaries of the tokens and will approximate the segmental phonological composition of the tokens. (Note: segmental phonological material consists of those sounds which follow one another in a linear sequence. Suprasegmental phonological material differs in that it is overlaid on segmental material and several suprasegmental phonological phenomena can occur simultaneously.) We say the transcription only approximates the actual phonological representation because we used a semi-phonemic system of notation only for those items in the speech sample which diverge from the pronunciation of the items as it would be in Standard American English. Thus, why if pronounced /wai/ as in Standard American English, was transcribed orthographically as why. However, if they (which is /ðei/ in a phonemic transcription of Standard American English) was pronounced as /dei/, with a /d/ substituted for the labial fricative /ð/, it was transcribed partially phonemically and partially orthographically as fry.

We decided to follow this practice of only partial phonemic transcription because of the time which complete phonemic transcription would demand. At this point of the analysis, pronunciations which diverged from those of Standard American English were of interest only in terms of several relevant indices.

In regard to the token, then, we made these particular distinctions in terms of our transcriptions:

(1) Only those tokens outside of games were counted.

(2) A token was defined as a meaningful segment composed of segmental phonemic material (as opposed to suprasegmental material such as stress, juncture, terminal contour); however, it was not necessarily transcribed phonemically nor need it have been in order to be identified.

(3) Phonological material counted as one token if it was separated from other phonological material by silence. This means we were not defining tokens as dictionary entries, since a token may be a composite of several lexical items such as did and you in the token (in phonemic transcription) /ǰuw/.

Type

The interest here was in the number of different types of tokens used by the subject. A type was defined as a phonological segment different from other segments in both phonological shape and meaning. Thus, different pronunciations of what was clearly intended as the same meaningful unit were counted as only one type. However, man and men both met the criteria set down as defining a type (i.e., they both represented changes in phonological shape correlated with change in meaning) and therefore were counted separately as individual types. It is important to note that only the first occurrence of a type was counted as part of the total number of types occurring in any segment of tokens.

Presumably, the ratio of types to number of tokens would measure two things: (1) vocabulary size (since the repeat of the same token does not increase the type count) and (2) the grasp of the rules governing English inflection. The vocabulary size would be indicated since the repeat of the same token does not increase the type count. The command of the rules of English grammar was shown in the size of the type total. For example, if the subject correctly used both singular and plural forms of nouns which have irregular plurals (such as man and men), the type-token ratios should have reflected this usage in that man and men would be counted as two different types as well as two different tokens. Since both forms might be present, yet wrongly used, (2) is applicable only if in both cases the forms are acceptable. Thus Index 10, unacceptable tokens, serves as a check on this index.

The ratio of types to tokens necessarily becomes smaller as the language sample becomes larger. Thus, a counting of types in relation to an uncontrolled number of tokens discriminates against the speech sample which is longer; the possibility that words (i.e., tokens) will be repeated becomes greater than the likelihood that new words will be introduced. (We must remember that only the introduction of new words results in an increase in the number of types.) In order to control this factor, we counted types only for the first 100-token segment of the speech sample.

The Token Count and Unacceptable Tokens

For each transcript we have a count of the tokens which have been uttered. This count was based on the tokens which occurred within the narrative sequence but not within the communication pauses. Of these tokens, some are termed "unacceptable" under four different categories on the list of Base Indices. All tokens, whether unacceptable or not, are counted in the token count. A list of the criteria for unacceptable tokens follows:

a. Tokens which are phonologically unacceptable

On the revised list of indices we considered under this heading the following phonologically unacceptable tokens:

1. The substitution of either /d/ or /t/ or /ʔ/ (where ʔ = glottal) for the anticipated Standard American English /ɾ/ (which is realized by a single touch of the tongue on the alveolar ridge). This substitution occurs in an inter-vocalic environment.

   Examples: Standard American English putting, in ortho-phonemic notation, putin, is replaced by /pudin/.
   Standard American English little in ortho-phonemic notation, litel, is replaced by lil.

2. The substitution of /ʔ/ for the final stop (which is normally an unreleased articulation in Standard American English) in a consonant cluster. This substitution occurs in a word final position.

   Example: Standard American English hand, in ortho-phonemic notation, hænd, is replaced by hæn.

   Note: We did not count those phonologically unacceptable tokens which represented a replacement of /n/ for /ŋ/, ortho-phonemic ng, in word final position, such as "talking" pronounced as /tɔkin/ rather than the standard /tɔkiŋ/.

3. The substitution of a voiced stop consonant in a consonant cluster by a homorganic voiced stop. This generally occurs in word final position.

   Example: Standard American English just is replaced by jus.
   (Note: The t is made in the same position in the mouth as s; therefore, it is called "homorganic.")

4. The substitution of /d/ for the first member of a consonant cluster.

Example: Standard American English that's in morphophonemic notation, [illegible], is replaced to [illegible].

b. Tokens which are grammatically unacceptable

Grammatically unacceptable tokens were defined as those tokens with morphological constructions differing from those present in Standard American English. Incorrect verb forms were not included.

Example: mens or mans for the Standard men
his'n for the Standard his

c. Tokens which are unacceptable according to usage practices

These tokens represent a departure from Standard American English practices of usage, although they may conform to the phonological and morphological rules of Standard American English.

Examples: "he ain't lookin'" for the Standard "he isn't lookin'"
"he was fixing to do that" for the Standard "he was going to do that"

d. Tokens which are unacceptable due to their non-English nature

Such tokens seemingly represent constructions of words which conform to the phonological and morphological rules of Standard American English, but which alter the meaning to such a degree that they are unintelligible as "American English" utterances.

Examples: "The [illegible] is there" for the Standard "The artist is there"
"The [illegible]..." for the Standard "The lady painting"

Narrative Segments

Content served as the basis for defining a narrative segment. Such a unit included that part of the speech flow dealing with a single subject with several subjects which were all elements in one "story sequence." Thus by "subject" we do not mean "sentence subject according to grammatical criteria," but rather "discussion from a single point of view centering around one notion." For example, in the following transcript segment, note where the narrative segment breaks are indicated:

There's a man coming down the ladder and uh there's a l-lady looking at 'em coming down through the uh--chute and uh one's got a ho- a horn and uh (this) other one's got a tube and uh the l-lady looks like looks very scared and uh (end of first narrative segment) she has a long a purse (end of second narrative segment.)

"She has a large purse" forms a separate narrative segment because even though "She has a large purse" is about "the lady" who is mentioned in the preceding sentence, this sentence represents a break from the "story sentence" of which "the lady" was a part. Indeed, "She has a large purse" is a kind of enumeration and not a story at all.

A subject would often begin a narrative segment with a "topic sentence" such as "I see people, trees, buildings--and cars, and there's three men and two ladies." He would then go on to comment specifically on some one object in the "topic sentence": "An there's windows in the building and there's a fence around around the sidewalk at the grass." In such a case, all of the material was counted as constituting one narrative segment. Of course, if what had followed the "topic sentence" had been a specific comment on objects not mentioned in the topic sentence, then more than one narrative segment was counted as present.

In the case of enumerations of items, each with few or no modifying segments following, the following convention was followed: we marked as one narrative segment any enumeration linked together by a cover phrase such as "I see .... 1, 2, and 3." However, an enumeration without a cover phrase was counted as a series of narrative segments. Thus, "In the picture there's trees, a lady painting a picture, two men watching...." would be one narrative segment, but "Trees, a lady painting a picture, two men watching..." would include three narrative segments. An enumeration without a cover phrase which occurred in response to a direct question, such as "What do you see?" was counted as one narrative segment.

One- or two-syllable responses to direct questions (yes, no, uh-huh, etc.) were included in the preceding or succeeding narrative segment, whichever made the most sense.

Communication Unit

A communication unit was defined by structural criteria as opposed to the content criteria which marked the narrative segment. However, because the phenomena which could be taken as structural criteria for determining communication units sometimes occurred in conflict with each other, we counted three different kinds of communication units.

Type #1 A speech segment was called a communication unit of this type if it was marked by a complete intonation pattern of a 2-3-1 relative pitch pattern and a terminal contour but did not consist of a grammatically independent clause on the segmental phonemic level. An example of such a communication unit is "on the building," when marked by a self-contained, complete intonation pattern.

Type #2 A speech segment was called a type #2 communication unit if it consisted of a grammatically defined independent clause on the segmental phonemic level (i.e., if a noun phrase and its verb phrase were present), but was unmarked by single, complete intonation pattern. Regardless how shown that examples which can be characterized as type F2 units usually occurred in the following patterns: one complete intonation pattern (therefore one type F1 communication unit) including two or more independent clauses (therefore two or more type F2 communication units). Thus, all speech material of type F2 units would also be the material of less than one, one, or more than one type F1 units. For example, consider the following speech segment:

> "Lady is wearin a black fur coat la- an the lady next to her is wearing a black coat and a black hat."

One complete intonation unit covered the entire speech segment; therefore, it was counted--as a whole--as one type F1 unit, *and* two type F2 units since it includes two complete independent clauses.

Type F3 A speech segment was termed a type F3 communication unit if a *complete intonation unit and a single complete independent clause occurred simultaneously*. For example, "The lady is lookin at the little man," if uttered in a single 2-3-1 intonation which terminates with a falling terminal contour, would be a typical type F3 unit. A simple sentence spoken in the Standard American English dialect would be classed a type F4 unit. Experience has shown that most utterances in the corpus of this study could be classified either as type F4 units, or as type F3 units followed by a type F1 unit (i.e., a modifying phrase which is spoken with a separate complete intonation unit).

Experience has also shown that it is sometimes difficult to decide for this corpus where, in fact, a completed intonation pattern occurred in such utterances as "The lady is paintin a picture of a man hanging." If two patterns are complete, one after "paintin" or "picture" and the other after "hangin," we must count "The lady is paintin (a picture) as one type F3 unit and "(a picture) of a man hangin" as one type F1 unit. However, if there is room for doubt whether, in fact, two complete intonation patterns exist, by convention we are biased toward counting type F3 units. Therefore, we would judge the entire utterance discussed above, "The lady is paintin a picture of a man hangin" as one type F3 unit.

We have had to exercise particular care in regard to two situations to avoid "over-counting" the number of type F1 units. The first is the occurrence of a pleonastic (or double) subject, as in "I see a man he paintin a picture." If the intonation contour clearly constituted only one complete pattern, the analyst disregarded the double subject ("man he") and counted the entire utterance as one type #3 unit. (It is, of course, possible that the subject could have marked off the above example into two separate intonation units: "I saw a man"// "he paints a picture"//, but the analyst took care to avoid being biased by the occurrence of the double subject.) Secondly, in her assessment of the intonation material, the analyst also took care to remember the necessity for a falling terminal contour to mark a single complete intonation unit. If a subject ended a possible intonation unit with a sustained terminal contour (signalling an incompleted communication unit), the analyst did not mark a completed intonation pattern, which would relegate that followed to a separate communication unit. Rhythm of speech also played a part in marking actually completed intonation units.

The decision to count these different kinds of communication units grew out of the corpus of speech samples and was not dictated by any _a priori_ definitions. In fact, _a priori_ considerations would lead us to expect type #3 communication units only, with the simultaneous occurrence of single complete intonation patterns with complete grammatically defined independent clauses. However, experience has shown that type #1 units occurred very often in conjunction with type #3 units in our corpus. We can attribute this occurrence to any one or all of a number of possibilities. That is, we can say the patterns found are possibly characteristic of (1) the speech of the educable mentally retarded, or (2) of the inner city resident, particularly if he is a child of recent non-white immigrants from the South, or (3) of any speaker of Standard American English in an impromptu situation. Further research would give us more information about the factors behind the production of the patterns we find in our corpus; for the present, however, it will suffice to note that supposedly atypical patterns do occur and to identify them accurately according to their characteristics.

Communication Mazes

Mazes were defined as incomplete fragments of thought such as false starts, incomplete words, and repetitions which add nothing to the content of the speech sample. The production of mazes resembles very much the behavior of a person trying to find his way in an actual spatial maze. The actions of a rat in a psychological lab trying to find his way out of a multiple T-maze offers a second analogy. For our purposes a maze was defined as an utterance without a functional semantic relation to the narrative segment in which it occurs. After a maze has been deleted from a narrative segment, the remaining utterance is a communicatively-viable unit in that further deletion would alter or negate the semantic content.

Repetitions were counted as mazes only under certain circumstances. If they added nothing further in content to the previous utterance, they were termed mazes. For example, in "I see the trees, the trees," the repeat of "the trees" was counted as a maze. But in "I see the trees, the beautiful trees," the repeat--because it included an amplification--was not so classified. Similarly, if the subject finished up his speech sample with "That's all. That's all," using the repeat for emphasis, this purposive iteration was not counted as a maze.

Unresolved Mazes

A subject may resort to one of two strategies when he gets into a maze:

a. He may *resolve* the maze. That is, he may go into a "blind alley" characterized by false starts, etc., but then move on to complete the thought sequence preceding entry into the maze. A maze is said to be resolved if:

   1. When repeated or altered to a correct form it becomes part of a communication unit.

      For example, in the following sentences, the portion enclosed in brackets represents a resolved maze:

      [S-] She was going home.

      [She] She was going home.

      [She was go-] She was going home.

   2. When dropped, it is followed by an utterance related to that which preceded the maze.

      For example:

      The lady is painting a picture. [And there's a-a]
      She's painting a picture of a man.

b. The maze may go *unresolved*. That is, the subject may enter into one or more blind alleys characterized by incomplete words, false starts, etc., and then proceed with utterances quite unrelated to the thought content of what preceded the maze. For example:

   The lady is painting a picture. [And there's a-a-.]
   [The man a-] I see buildings behind and....

In general, a coherent narrative segment should remain when all the mazes are removed. The mazes are not considered part of any communication unit; in fact they typically interrupt a communication unit both in regard to the completion of an intonation pattern and the syntactic ordering of grammatical units.

Finite Verbs

Finite verbs were defined as those which stand as the main verb phrase in any independent clause. They were thus distinguished from participles, which stand as modifiers of noun phrases, and from gerunds and infinitives, which stand as noun phrases. Within any independent clause, a finite verb phrase is required; but, in our corpus, required verbs were often absent. In the counting of number of finite verbs, however, we counted all verbs which were required, *whether they were actually present or not*. A participle preceded by a form of "to be" standing as an auxiliary was counted as one verb (e.g., "is going").

Table 1

Speech Sample Derived Indices

Key to IBM Card Fields with Description and Formulas

IBM Card Columns	Derived Index Number	Description	Formula (In terms of raw index numbers)
First Card			
1-4		Code Number	
5		Cell Number	
6		Test Number	
7		Form Code	
8-10		Blank	
11-13	1	Words per minute	5/1
14-20	2	Mean words per successive sentence	5/2
21-24	3	Ratio of syllables in pauses to total syllables in transcript	13/11
25-28	4	Ratio of type to tokens	6/(5**)
29-32	5	Ratio of acceptable tokens to total tokens	[5 - (8 + 9 + 10)]
33-36	6	Ratio of unacceptable phonological tokens to total tokens	7/5
37-40	7	Ratio of unacceptable morphological tokens to total tokens	8/5
41-44	8	Ratio of unacceptable usage tokens to total tokens	9/5
45-48	9	Ratio of unacceptable English word tokens to total tokens	10/5
49-52	10	Ratio of unresolved pauses to total pauses	4/3
53-56	11	Ratio of incorrect verb forms (excluding omission of "to be") to total finite verbs (13 of raw indices)	14/13
57-60	12	Ratio of incorrect verb forms which are omissions of "to be" with main verb following to total finite verbs (13 of raw indices)	15/13
61-64	13	Ratio of incorrect verb forms which are omissions of "to be" as main verb (no verb in sentence) to total finite verbs (13 of raw indices)	16/13
65-68	14	Ratio of number of type 1 communication units to total number of tokens	17/5
69-72	15	Ratio of number of type 2 communication units to total number of tokens	18/5
73-76	16	Ratio of number of type 3 communication units to total number of tokens	19/5
77-80	17	Average number of syllables per token	(11 - 12)/5

Second Card

1-12		Same as cols. 1-12 on first card (above)	
13-16	18	Ratio of unconventionally spoken complete intonation patterns to total number of complete intonation patterns	17/(17 + 19)
17-20	19	Ratio of unconventionally spoken independent clauses to total number of independent clauses	18/(18 + 19)

Table 2

Speech Sample Raw Indices

Key to IBM Card P1-14a with Descriptions

IBM Card Columns	Raw Index Number	Description
1 - 4		Code Number
5		Cell Number
6		Test Number
7		ITPA Code
8 - 17		Blank
18 - 20		Type 3 communication unit mean level of complexity
21 - 23		Number of well-formed type 3 communication units
24 - 26	1	Time in minutes and five second intervals
27 - 29	2	Number of narrative segments
30 - 32	3	Number of communication mazes
33 - 35	4	Number of unresolved mazes
36 - 38	5	Total number of tokens in communication units
39 - 41	6	Total number of types in communication units
42 - 44	7	Unacceptable selected phonological tokens
45 - 47	8	Unacceptable morphological tokens
48 - 50	9	Unacceptable usage tokens
51 - 53	10	Unacceptable English word tokens
54 - 56	11	Number of syllables in transcript
57 - 59	12	Number of syllables in mazes
60 - 62	13	Number of finite verbs, including instances where a verb should be but was omitted
63 - 65	14	Number of incorrect verb forms (including omission of "to be")
66 - 68	15	Number of omissions of "to be" with main verb following
69 - 71	16	Number of omissions of "to be" when required (i.e., no verb in sentence)
72 - 74	17	Number of type 1 communication units
75 - 77	18	Number of type 2 communication units
78 - 80	19	Number of type 3 communication units

From this total count of required finite verbs, we were interested in obtaining the following additional information:

a. Number of incorrect verb forms including wrong tense, wrong inflection for person, or omission of the verb altogether. For example:

The sky white (is is omitted as a main verb)
That man lookin. (is is omitted as an auxiliary)
The lady have a purse. (have is the incorrect form
of the verb)
The little man down from the box. (main verb omitted)

b. Number of times "to be" is omitted with a participle following (e.g., "The lookin at the man.")

c. Number of times "to be" is omitted when it is required as a main verb in the clause. (e.g., "The sky white.")

These raw indices were punched on IBM cards in the format shown in Table 1. A program was written using the formulas shown in Table 2 and the derived indices so obtained were punched out in the format shown in Table 3.

B. Interpretation of Derived Language Indices

It was assumed that the instructional endeavors occurring during the academic semesters would have a positive effect on the language encoding of the pupils as indexed by the derived indices. The derived indices consisted mainly of ratios which allowed a comparison of transcripts which were of varying length and duration. In Table 3

Insert Table 3 about here

---

are shown groupings of the 19 indices along with the directional orientation of the measures. Thus, a "+" indicates that an increase in magnitude is associated with a desirable direction of change while a "-" indicates an increase is associated with an undesirable direction of change.

Verbal Output

1. Words per minute - index of the word productivity of the pupil in terms of speech rate.

3. Mean words per narrative segment - may be interpreted as an index of language maturity.

4. Type-token ratio (TTR) -- indicates the degree of variability or flexibility of language usage and has been found to be greater in more intelligent children.

17. Average numbers of syllables per token (word) -- it is assumed the capacity to use words of more than one syllable reflects a goal of education. Although this index has been found to correlate highly with TTR, for our population the TTR may be quite high, yet this index might be quite low.

Table 3

Grouping of 19 Derived Indices and Desirability of Increased Magnitude of Scores

Group	Index # From Tables 1 and 2	Desirability Direction
Verbal Output	1	+
	2	+
	4	+
	17	+
Pauses	3	-
	10	-
Acceptability Measures	5	+
	6	-
	7	-
	8	-
	9	-
	11	-
	12	-
	13	-
Intonation Patterns	14	+
	15	+
	16	+
	18	+
	19	+

+ Increase in score associated with desirable change

- Increase in score associated with undesirable change

Mazes

9. Ratio of syllables in mazes to total syllables in transcript -- may be considered an index of the adequacy of communication efforts, since it is obvious that little, if anything, is communicated by the individual who spends most of his time in mazes. (It should be noted, however, that there is no guarantee that the subject who has no mazes is communicating effectively.)

10. Proportion of unresolved mazes -- is used to tap cohesiveness of thought processes. It is assumed that a greater proportion of unresolved mazes indicates less cohesiveness of thought.

Acceptability Measures

5. Proportion of "acceptable" words -- the number of acceptable words (as defined) divided by the total number of tokens in the transcript outside of mazes. In using this index, we assume that correct pronunciation is a goal of education. Indices 6, 7, 8, and 9 provide a more detailed attempt to tap the nature of "unacceptable" words.

11. Proportion of verb forms incorrect -- measures one part of the grammatical aspect of the educational endeavor. This index plus the two following were demonstrated by Loban (1966) to vary markedly as a function of grade level.

12. Proportion of incorrect verb forms which are omissions of "to be" with main verb following.

13. Percentage of incorrect verb forms which are omissions of "to be" i.e., no verb in sentence.

APPENDIX B

Standardized Procedure for Use of Addams' Drawings to Elicit Representative Verbal Responses Over a Three-Minute Time Interval

The purpose of this phase of the testing is to obtain a three-minute representative sample of the subject's speech. The interest is not in the thematic content. If the examiner has had previous experience with recorders in the type of situation described here, it is suggested that whatever physical arrangement of the equipment which has worked well in the past be used. Otherwise it is suggested that the recorder be placed within easy reach of the examiner and that the microphone be placed between one and two feet from the subject and facing him. Set volume and tone controls at middle position.

The stimuli used to elicit the speech production are two Charles Addams' drawings. These should be placed face down on the table when not in use.

Procedure

Turn the recorder on and when warmed up indicate in a clear voice the name of the subject, his birth date and your name as the examiner, i.e., "John Doe, born 3-14-52, examiner, Harriet Smith." Stop the recorder and give the subject the following instructions:

> We are going to play a game in which I want you to tell me a story about some pictures I will show you. The tape recorder will be on while you are telling the story, so I would like you to speak clearly. [Show the child one of the drawings and continue the instructions]. You might tell me who is in the picture, what is happening in the picture, or what the people are doing. You might also tell me what you think happened before and what will happen. O.K., go ahead.

Turn the recorder on.

Since the purpose is to obtain three minutes of speech production, let the child continue talking for this length of time. At three minutes, use a natural break and say: "That was fine. That was a good story." Turn the recorder off.

If the child does not respond to the first drawing within about 15 seconds, or if the child stops responding for 15 seconds, use a prompt of the followin nature:

> "Tell me about the picture"
>
> or
>
> "Tell me more about the picture"

Allow 15 seconds to elapse; and if the child has not started to speak, use another prompt as before. If he still hasn't responded to the picture after 15 seconds, try the second picture and go through the same procedure. *Do not use more than two prompts per picture* and space them about 15 seconds apart.

Our intent is to analyze the stories told by the subjects in terms of usual encoding indices such as length of the verbalizations or total word count, the average sentence length, and the mean of the five longest remarks in each verbalization, type-token ratio, etc. Thus, if the subject responds well to the first picture, there is no need to use the second one.

Points in speech sampling procedure to which particular attention should be given:

To reiterate, the aim of this procedure is to elicit (obtain) and record at least three minutes of actual speaking by the subject; not to merely consume approximately three minutes of recorder or tape time with each subject. Hence, it might be necessary to consume five or six, or even more, minutes of recorder time to get these three minutes of actual speaking by the subject. Of course exceptions to this standard should be made where one is clearly and simply not going to get this much speech from a subject within a *reasonable* length of time.

Subjects should *actually be permitted* pauses of at least 15 seconds duration *without* any prompting, before a prompt *is* used; and a maximum of only two prompts per picture should be used.

Since the aim of the procedure is to obtain three minutes of actual speaking by the subject, if this is accomplished (within the standards for use of prompts described above) by use of only one picture, the second picture need not be used. However, if this much actual speaking is not obtained by use of only one picture, then the second picture should certainly be used.

8. Factors Influencing Language Indices

One study (Cowan, et al., 1967) has investigated a number of variables which might influence mean length of response (MLR), i.e., the average number of words per remark emitted by Ss (5 through 11 years of age) in responding to 10 different pictures. They found that the stimulus influenced the length of response, that age, sex and socio-economic status effects interacted, and that the experimenter effect interacted with the age and sex of the subjects. Their findings suggested that the measures we are obtaining might be quite unreliable. Accordingly, several substudies were undertaken to answer the following questions as they apply to the present project:

1) Does the stimulus, the testers or the repeated occurrence of testing influence 10 indices of language functioning used in the present project?

2) What is the test-retest reliability of 10 indices when subjects are retested within the same week?

3) What is the reliability of two analysts who are called upon to analyze transcripts independent of one another?

4) What are the intercorrelations among the 10 indices on three occasions when the analysis of the transcripts represents agreement between two analysts?

Procedure (Question 1)

Thirty-two primary and intermediate [illegible] pupils attending summer school were used in the study. These children were considered representative of the population from which the project drew its sample with one important difference. They were having sufficient difficulty in school to seem to warrant the additional attention during the summer. Thus, they were probably functioning at a lower level on the average than the project sample as a whole even though some of the project children were also picked up in the study sample.

A Latin square design was used to evaluate the effects of the independent variables. The design was set up as follows:

	Time 1	Time 2
Sequence 1	Form A	Form B
Sequence 2	Form B	Form A

Time refers to the test and retest given within the week. Sequence 1 refers to the fact that different testers were used and sequence 2 indicates that the same testers were used. Form A includes a set of ten Charles Addams drawings which were used as stimuli. Form B includes a different set of the same kind of stimuli.

The procedure used in eliciting the language sample is described elsewhere (Addendum A) as are the raw indices which were obtained from the transcribed samples. The derived indices, corrected for time, etc., which formed the dependent variables in the analyses are also described on page 14 in the previous section.

Subjects were assigned randomly to the two sequence conditions and the retesting of the subjects varied from two to five days.

Results and Interpretation

The analysis of variance tables for the 10 derived indices are shown in Table 1 (a) through (j). In these analyses the group or sequence effect can be considered as part of the form x time interaction. With 1 and 30 df, the F ratio required at the .05 level is 4.17 and at the .01 level, 7.56.

Insert Table 1 about here

A summary of the findings of these tables is:

1) The Time variable did not have an effect on any of the 10 indices. That is, there was no significant change in the level of response of the group as a whole upon retesting within a week.

2) The sequence or tester effect was evident in two indices (Table 1, (d) and (f)).

(d) Type-token ratio. The first sequence consisted of using two different testers at Time 1 and 2 whereas the second sequence used the same two testers both times. This particular sequencing was not intended initially, but illness of one of the testers forced us to go this route. As it stands however, sequence 1 resulted in a TTR of 53% while sequence 2 resulted in a TTR of 61.7%. This difference can be interpreted as suggesting that the use of the same testers on test and retest results in a greater variety in verbal output as indexed by the TTR. In retrospect, a shortcoming of this index as measured, lies in the fact that it was not computed as a function of some specified number of tokens. For, it does seem that the more words the individual produces, then overall the number of different types will decrease. The shortness of the transcripts (most less than one page double spaced) and the use of two cards may have, however, reduced the relevance of the above criticism.

(f) Unresolved mazes as percent of total number of mazes. Using different testers resulted in an average of 15.3% of the mazes being unresolved, while the same testers resulted in 31% of the mazes being unresolved. Our interpretation of this measure is that it indexes cohesiveness of thought processes. That is, the higher the percentages, the less coherentness present. Using this reasoning, we have to assume that using different testers within a week's time results in greater coherentness than does using the same testers. Interpretatively this could mean that seeing the same tester within a few days makes the subject attempt to express ideas and use syntactic structures not readily available in his repertoire, hence more unresolved mazes.

3) The form effect was evident on two indices also (Table 1 (a) and (f)):

(a) Words per minute. Form A resulted in an output of 48.4 words/minute while Form B resulted in 57.3 words/minute. These findings suggest that if one wants to maximize verbal output then the two drawings comprising Form B are to be preferred.

(f) Unresolved mazes as a percent of total number of mazes. The percentages here were 20.9% for Form A and 17.4% for Form B. Form B can then be said to provide the possibility of more cohesive thought processes in verbalizing about the drawings.

Conclusions

On the basis of the analysis of this substudy, the following conclusions appear tentatively warranted.

1) Level differences in 10 language indices were not evident when retesting a group of 33 EMR children using two different sets of Charles Addams drawings and different testers. This suggests that the instrument as such should have value in repeated testing for the evaluation of different treatments taking place during the interval. This assumes that all pertinent tester and stimuli variables can be controlled.

2) The interpretation of the effect due to use of the same or different testers on retesting puts one in the position of Buridan's ass. This is assuming that we are concerned about having the testing situation set up in such a way as to maximize the possibility of tapping to the fullest extent the language encoding capability of the child. For, using different testers, as compared to the same ones, results in lessened complexity of output as indexed by the TTR. On the other hand, using different testers also resulted in greater coherentness of thought processes (i.e., fewer unresolved mazes). It could be of course that inherent in the relationship is the fact that lessened complexity does indeed covary with greater coherentness of thought process. That is, fewer new forms are tried, hence fewer different types, but also fewer unresolved mazes. Thus, having different testers on two occasions may be cause for a more restrictive functioning which manifests itself in the directions the two indices took.

3) The Form B set of drawings would seem to be preferable to the Form A. Form B resulted in greater complexity of verbal output and at the same time resulted in fewer unresolved mazes.

Procedure (Question 2)

The speech samples of the 33 subjects obtained in the study were transcribed using two transcribers to check on one another. Analysis was done in terms of the raw indices described elsewhere in the report. Three analysts reached agreement as to the scoring of the transcripts. The derived indices were then calculated. The test-retest reliabilities of the indices were computed using the variance information contained in the source tables shown in Table 1. The formula used was:

$$r_x = 1 - \frac{\text{MS error}}{\text{MS subjects}}$$

In effect, this reliability estimate is independent of the effects of forms, sequences and times of administration and as such may be considered to be approaching the "true" measure of reliability of the various indices.

Results

Table 2 column (b) shows the test-retest reliabilities of the 10 indices.

Insert Table 2 about here

All but two measures had significant and frequently quite high reliabilities. These two exceptions were (Table 2) (3) *"Acceptable" words as percent of total uttered* in narrative segments and (9) *Percent of incorrect verb forms which are omissions of "to be"*. Both of these indices have been modified and clarified as to meaning. (See attached analysis and interpretive formats which is a revision of the one on which these data were based.)

Procedure and Results (Question 3)

The concern of this section was to examine analyst reliability with a view to being able to pinpoint those measures which required further clarification. Two of the trained analysts were given 10 transcripts from the main project to analyze independently. Intercorrelations were then run with no correction for attenuation. Table 2, column (a) shows the results of this analysis. Three measures stand out as in need of clarification: (7) Mean words per minute, (8) Percentage of incorrect verb forms which are omission of "to be" with main verb following and (9) Percentage of incorrect verb forms which are omission of "to be": no verb in sentence. Index (3) syllables in mazes as percent of total syllables in transcript has proved to be most difficult to work with, but because of its high test-retest reliability (col. b) and logical relevance to educational endeavor, the effort would appear to be worth it.

Procedure and Results (Question 4)

The intercorrelation of the various measures was sought. The 33 subjects tested during the summer were used for this purpose. Agreement was obtained between three analysts prior to intercorrelating the 13 indices. The intercorrelations are shown in Table 3. These intercorrelations must be considered

Insert Table 3 about here

of limited value at the moment because of the modifications which have taken place in the indices. The fairly frequent and relatively high correlations suggest that the variance associated with the indices could be meaningfully reduced through a factor analysis using in addition the various achievement and linguistic scores obtained in the project.

Table 1

Source tables for analysis of variance of the 10 dependent measures

a) Words per minute

Source	SS	df	MS	F
Sequence	4,254.3	1	4,254.3	3.70
Ss/Sequence	34,515.2	30	1,150.5	--
Time	424.3	1	424.3	2.63
Form	1,278.0	1	1,278.0	7.92*
Error	4,842.5	30	161.4	--
	45,314.3			

b) Mean words per narrative segment

Source	SS	df	MS	F
Sequence	115.8	1	115.8	<1
Ss/Sequence	12,063.9	30	402.1	--
Time	16.7	1	16.7	<1
Form	7.8	1	7.8	<1
Error	681.0	30	22.7	--
	12,885.3			

c) Syllables in pauses as a percent of total syllables in transcript

Source	SS	df	MS	F
Sequence	4.1	1	4.1	<1
Ss/Sequence	3,270.6	30	109.0	--
Time	95.8	1	95.8	3.11
Form	41.4	1	41.4	1.34
Error	923.1	30	30.8	--
	4,337.2			

d) Type-token ratio

Source	SS	df	MS	F
Sequence	1,011.0	1	1,011.0	6.69*
Ss/Sequence	4,526.9	30	150.9	--
Time	105.0	1	105.0	1.51
Form	25.7	1	25.7	<1
Error	2,083.1	30	69.4	--
	7,751.7			

e) "Acceptable" words as percent of total uttered in narrative segment

Source	SS	df	MS	F
Sequence	.1	1	.1	<1
Ss/Sequence	53.9	30	1.8	--
Time	5.2	1	5.2	2.6
Form	4.3	1	4.3	2.2
Error	60.4	30	2.0	--
	123.7			

f) Unresolved mazes as percent of total number of mazes

Source	SS	df	MS	F
Sequence	3,918.8	1	3,918.8	4.37*
Ss/Sequence	26,910.1	30	897.0	--
Time	124.9	1	124.9	<1
Form	2,141.4	1	2,141.4	4.36*
Error	14,751.0	30	491.7	--
	47,846.2			

g) Percentage of verb forms incorrect

Source	SS	df	MS	F
Sequence	69.3	1	69.3	<1
Ss/Sequence	29,898.6	30	996.6	--
Time	33.3	1	33.3	<1
Form	1,438.3	1	1,438.3	4.01
Error	10,758.1	30	358.6	--
	42,197.6			

h) Percentage of incorrect verb forms which are omission of "to be" with main verb following

Source	SS	df	MS	F
Sequence	4.1	1	4.1	<1
Ss/Sequence	21,866.5	30	728.9	--
Time	41.5	1	41.5	<1
Form	98.3	1	98.3	<1
Error	8,035.2	30	267.8	--
	30,045.3			

i) Percentage of incorrect verb forms which are omissions of "to be": no verb in sentence

Source	SS	df	MS	F
Sequence	3.3	1	3.3	<1
Ss/Sequence	951.3	30	31.8	--
Time	1.5	1	1.5	<1
Form	14.5	1	14.5	<1
Error	811.6	30	27.0	--
	1,782.2			

j) Average number of syllables per token

Source	SS	df	MS	F
Sequence	.03	1	.030	3.00
Ss/Sequence	.36	30	.012	--
Time	.00	1	.000	0
Form	.00	1	.000	0
Error	.15	30	.005	--
	.54			

Table 2

Intercorrelations of a) two analysts' analysis of the same transcripts, and b) test-retest of the instrument with scores based on agreement between three analysts

Index	a) Analyst Reliability	b) Instrument Reliability
1. Words per minute	.99*	.86*
2. Mean words per narrative segment	.58	.94*
3. Syllables in mazes as a percent of total syllables in transcript	.87*	.72*
4. Type-token ratio	.96*	.62*
5. "Acceptable" words as percent of total uttered in narrative segment	.69*	-.11
6. Unresolved mazes as percent of total number of mazes	.80*	.43*
7. Percentage of verb forms incorrect	.88*	.64*
8. Percentage of incorrect verb forms which are omission of "to be" with main verb following	.61	.43*
9. Percentage of incorrect verb forms which are omission of "to be": no verb in sentence	.42	.08
10. Average number of syllables per token	.97*	.50*
N	10	33

*Significant at the .05 level of confidence

Table 5

Intercorrelations between 10 derived language indices when analysis based on agreement between three analysts. (N = 33 subjects)

Scale (Time 1)	1	2	3	4	5	6	7	8	9	10
1. Words per minute	—	.46*	-.19	-.71*	.12	-.23	-.36*	-.35*	-.27	-.34*
2. Mean words per narrative segment		—	-.03	-.24	.08	-.14	-.37*	-.23	-.23	-.11
3. Syllables in nouns as a percent of total syllables in transcript			—	-.06	-.17	.31	-.02	-.21	-.17	-.29
4. Type-token ratio				—	.02	.11	.63*	.73*	.13	.77*
5. "Acceptable" words as percent of total uttered in narrative segment					—	-.06	.01	-.03	.03	-.01
6. Unresolved nouns as percent of total number of nouns						—	-.08	-.03	-.02	-.16
7. Percentage of verb forms incorrect							—	.84*	.23	.64
8. Percentage of incorrect verb forms which are omission of "to be" with main verb following								—	.21	.77*
9. Percentage of incorrect verb forms which are omission of "to be": no verb in sentence									—	.37
10. Average number of syllables per token										—

*Significant at the .05 level of confidence

3. Changes in Language Patterns as a Function of Race, Sex, and IQ

The major portion of this section was written by Richard Turner, presently at the Department of Special Education, University of Indiana. Reference to significance or non-significance of results are based on statistical analyses done on the data.

The data presented in Figures 1-4 can all be subsumed under the heading, "verbal output." Each of these measures is a quantification index which presumes that language can somehow be segmented into discrete units, and that these units can be counted or otherwise manipulated in a meaningful way. Such segmentation requires very restrictive definitions of each unit; without restrictive definitions ambiguity inevitably results.

The first figure (Fig. 1) presents the data for words per minute

---

Insert Fig. 1 about here

This measure counts only those tokens which are left over after the mazes are removed from the transcription. The definition of "token" used in the present study ("stretches of sound between silences") often requires that combinations of "words", as they are traditionally counted, be counted as one token. For example, one common traditional definition of "word" (though rarely made explicit) appears to be "stretches of print between white spaces." Under this definition, most studies simply have a secretary transcribe the tapes, and then the "words" are counted. Both of these restrictions, i.e., counting only the number of words outside of mazes, and the restricted definition of "token" probably have made the "token count" far different from any "word count" that appears in the psychological or speech pathology literature to date. No studies are available to which the present data can be directly compared. Therefore, all of the "count" data is unique.

It is questionable whether rate of speaking (Fig. 1) is meaningful in a linguistic sense. It is probably more meaningful in a sociological framework since rate of speaking appears to be quite dependent on social class background. Bernstein (1964) and others in his group (Lawton, 1963, 1964; Robinson, 1965) have presented a great deal of evidence in support of the hypothesis that language is used for different purposes by differing social class groups. Bernstein's theory states that the lower classes are likely to use language to manipulate ideas and notions. In maintaining the solidarity, *what* is said does not seem to be as important as *how* it is said among the lower classes. That is, the lexical and structural components of lower-class speech tend to be highly predictable, and paralinguistic features carry a heavy load in message transmission.

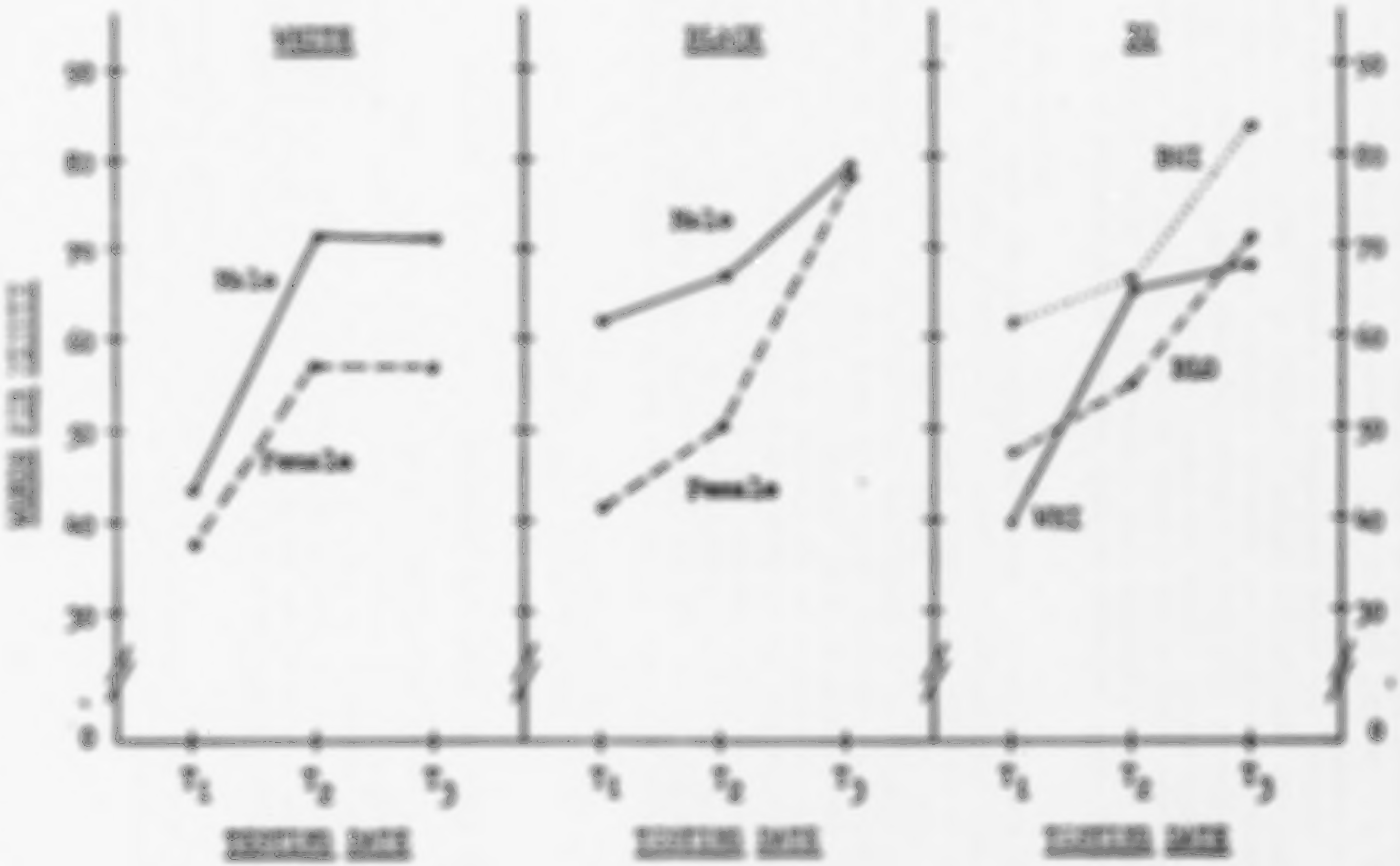


Fig. 1 -- Words per minute as a function of sex, race, and IQ level for [illegible] children over three testing dates.

Consequently, the verbal planning requirements of lower class speech tend to be less than those of the middle class, and the result is more fluent, and structurally predictable speech than is true for the middle classes.[5] This enables the lower classes to use more rapid speech because there are fewer pauses, hesitations and the like. Since both white and black groups used in the present study were from the lower classes, no such difference in rapidity of speech would be expected, and indeed, none such was found. Bernstein's theory also states that SR is not important in the verbal planning requirements of class language, and the SR data presented in Fig. 1 appears to be consistent with the theory.

In spite of the lack of significant differences, the patterns evident in the graphs are interesting in that they may be indicative of variables amenable to research in the future. If rate of speaking is a cultural rather than a linguistic index, there may be several factors which could be separated in future studies which use this index.

The first factor, of course, would be Lower Class vs. Middle Class, since no data for middle class children were obtained in the present study. Under Bernstein's theory the middle class children would be expected to speak at a slower rate than the lower class children. However, accurate interpretation of any data obtained would have to take into consideration both the amount of time and the number of pauses in speech. Data will be presented below to indicate that there are not many differences between the groups used in this study in the pause indices, but they could well appear in cross-SES class studies.

The second factor would be Race X Social Class. The patterns in Fig. 1 show that the white youngsters appear to have "peaked" their rate of speaking at T_2 while the black youngsters continued to increase to T_3 (and could well increase beyond the point which was actually found had more data been obtained). This phenomenon could be due either to a progressive "loosening-up" effect as the children continued to come back to the same experimental session, or to more rapid speech as the children grew older and more knowledgeable. The first interpretation seems the more likely, since there are few, if any, effects which can be explained on the basis of learning which took place during the test period.

The third factor concerns a possible Sex X Race interaction, which might also be affected by social class. Males and females play different roles in American societies, and Shuy, Wolfram, and Riley's (1967) data indicate that Social Class X Sex correlations appear in linguistic data. The data in Fig. 1 indicate different rates of speaking over time for Race X Sex. The white males and females at T_1 spoke at about the same rate, while the black males and females were quite far apart. By T_3 the black females had closed the gap, while the white males were speaking more quickly than the white females. It also could be that different rates of [illegible] could appear among white and black males and females depending on whether the situation is familiar or unfamiliar. If there is such an effect, it too could be affected by social class.

---

[5] Bernstein's measures of fluency and structural predictability are rather different from the measures employed in the present study, with the exception of the simple timed measure (words per minute). As a consequence, complete application of his measures is unnecessary, and the interested reader is referred to Bernstein's articles for elaboration.

Figure 2 presents the data on the mean number of words per narrative

---

Insert Fig. 2 about here

---

segment. This measure was designed as an index of the amount of talking each group did in relation to individual topics. Because the data on the white groups appears to be incomplete, final write-up will await checking of the data. Nevertheless, a few preliminary comments can be made.

This data cannot be affected by mazes since it concerns only those parts of the transcripts that are left over after the mazes are removed. It is affected by the definition of "token," however, because dialect features in either the speech of either blacks or whites could affect considerably the number of "tokens" to be counted. For example, if one dialect, but not the other, consistently uses tokens like "wanna" instead of "want to," the token count would be different for the two groups. The same factor can affect all of the measures which count tokens and should be borne in mind in the interpretation of the data.

The data on the blacks consistently dips at T_2. This time coincides with the racial riots in Milwaukee, and much of the data in the study is likely to be affected by the emotional atmosphere at the time. Fewer words per narrative segments might indicate taciturnity on the part of the Ss. Such taciturnity could well be expected from the black children in riot-torn ghettos. One interesting thing appearing in Fig. 2 is that the lower IQ black retardates recover far more by T_3 than do the higher IQ black retardates. The question of whether the lower IQ black retardates left the negative emotional atmosphere behind them more quickly than the higher black retardates must surely be raised. This interpretation, however, while suggested, is not completely clear from the present data because fewer words per narrative segment could well be the result of a deficiency in either the linguistic or cognitive process. The dip at T_2 certainly makes the taciturnity hypothesis an interesting one to consider, however.

Figure 3 presents the data for the mean number of syllables per token.

Insert Fig. 3 about here

If comparisons to other studies are desired, this measure might be affected by the removal of mazes from the transcript. The restricted definition of "token" undoubtedly affects the measure in its present form, but how the effect appears in the present data is unknown.

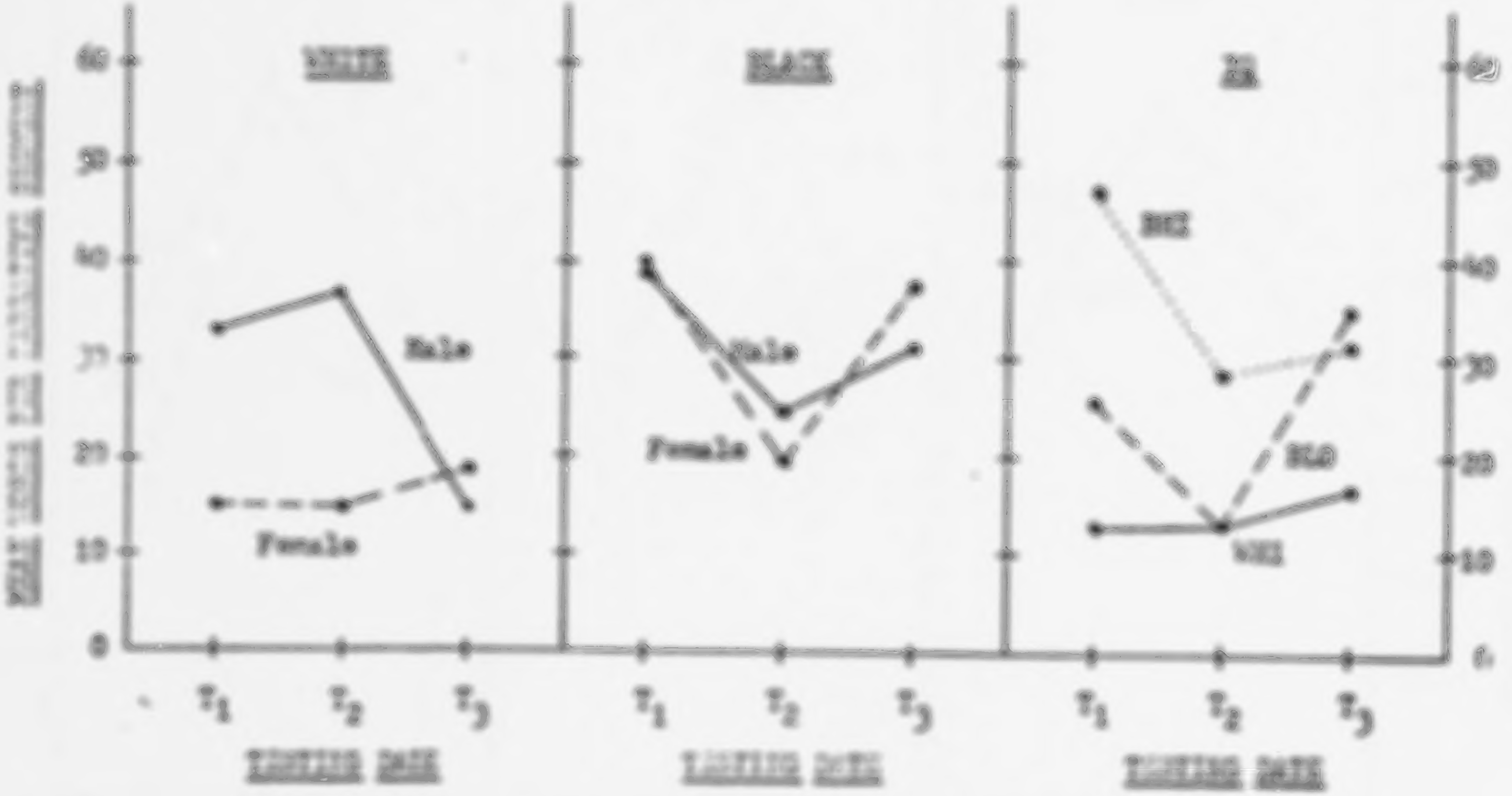


Fig. 2 — Mean words per narrative segment as a function of sex, race, and IQ level for EMR children over three testing dates.

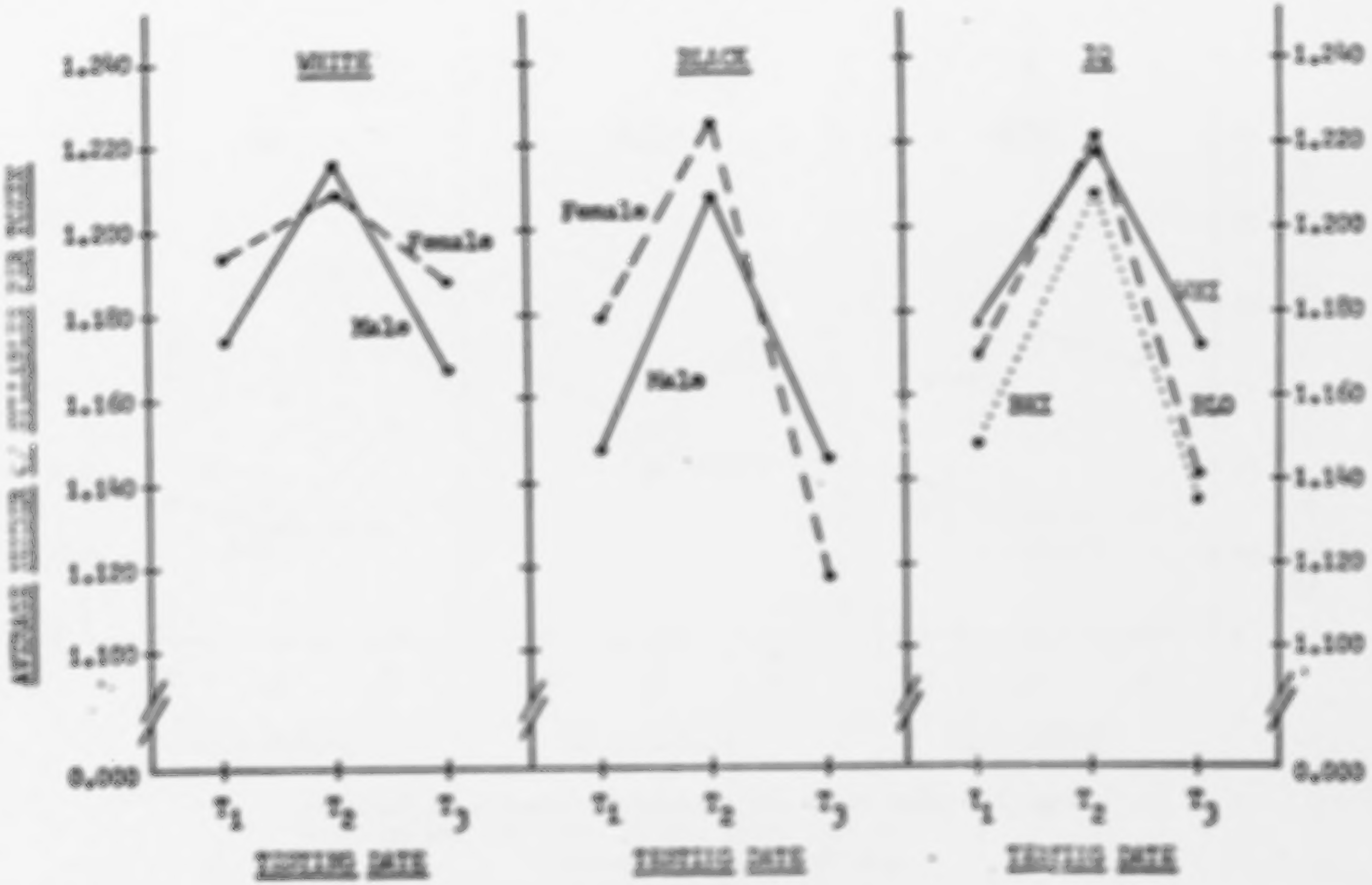


Fig. 3 — Average number of syllables per token as a function of sex, race, and IQ level for EBR children over three testing dates.

Significant differences appeared at T_1 and T_2 in the [illegible] data, and at T_3 in the [illegible] data. However, it is impossible to tell whether one group used longer lexical forms, or whether the values appeared to be long because of the use of "run-on" forms, e.g., "wanna."

The usefulness of this type of measure is unclear, even if the data were interpretable. That is, it is not clear whether the use of polysyllabic words would indicate a more sophisticated level of language development, or whether uni-syllabic words indicate a greater use of function words, and therefore, a higher level of linguistic functioning. The first interpretation would be supported by the correlation between IQ type vocabulary subscales and ability to achieve academically, and the second would be supported by the need for function words in such complex patterns as subordination, possessivization, etc. The question could be argued either way, and until the data are subjected to more complete analysis, it would have to remain open.

Figure 4 presents the data on Type-Token-Ratio (TTR). This measure is supposed to be an index of range of vocabulary, i.e., the number of different words used expressed as a ratio of the total number of words used. There was only one significant difference found between groups. Since it was at T_3, the time of the Milwaukee riots, its importance must be discounted.

Insert Fig. 4 about here

This data can be compared to the data presented by Loban (1963) since he, too, used the method of removing all mazes from the data before the counts were taken. The definition of "token" is unlikely to affect the comparison. For all groups the TTR is on the order of 20-30% lower than was found among Loban's low language-proficiency groups in all grades, K-3. Since the TTR was able to discriminate slightly between the high and low language proficiency groups in the Loban data, it is possible that the obtained TTR in the present study reflects a very low level of language proficiency for all Ss. However, it is also possible that the Charles Addams cartoons used as eliciting stimuli were of such a nature that the possible obtainable vocabulary was severely restricted. The TTR's are so much lower than they were in Loban's data that the latter interpretation is a distinct possibility.

In summary, then, none of the first four measures considered in Section I can be considered to be a good index of either linguistic development or linguistic proficiency. The rate of speaking measure appears to be more connected to cultural variants than to either IQ or sophistication. Mean words per narrative segment appears to be too susceptible to external events to be interpretable. Syllables per token is simply uninterpretable in the present context, and the TTR may have been artificially depressed by the eliciting stimuli, and is therefore also uninterpretable. The next two figures (Fig. 5 & 6) may have more relevance.

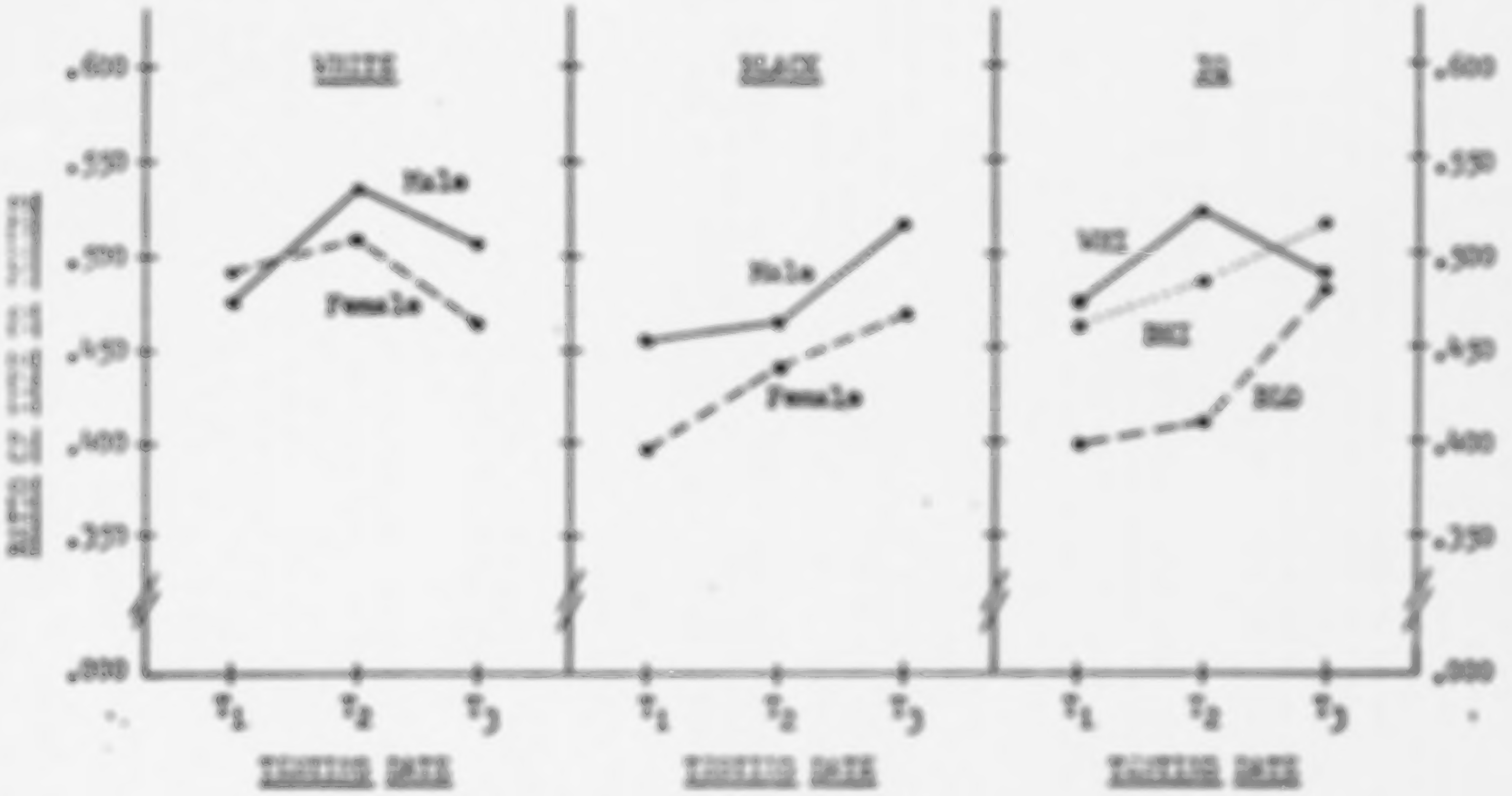


Fig. 4 — Ratio of type to tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

Figs. 5 & 6 present the data on mazes. Fig. 5 reflects the amount of speech spent in mazes, and Fig. 6 reflects the relative number of unresolved

---

Insert Figs. 5 & 6 about here

---

mazes. Again, the only significance is found at T_2, and then only for females. This must be discounted, and the conclusion must be drawn that all groups spend about an equal amount of their speech in mazes.

Loban's (1963) data indicate that his low language-proficiency group spent approximately the same amount of transcript time in mazes (.12-.09 words/maze in grades 2-6). In spite of the fact that Loban's data utilize words rather than syllables, the ratios are roughly equivalent to those found in the present study. Since Loban's low language-proficiency group had Kuhlman IQ's which were higher than the IQ's in the present data (99.3 mean IQ in grades 2, 4, 5, and 6, personal communication from Loban). It is also likely to yield significant differences when EMR's are compared to normals. The present data compared high IQ EMR's to low IQ EMR's, blacks to whites, and males to females. Since the children in this study appear to approximate (roughly) the low language-proficiency children in the Loban study in transcript-time spent in mazes, the question must be raised as to whether this measure can discriminate one level of "lowness" from another. This question, of course, assumes that the low IQ EMR's have less linguistic proficiency than the high IQ EMR's. There has been no data presented yet which conclusively demonstrates the validity of this assumption, either in the present study or in any other study (Semmel, 1968). The state of knowledge is still very poor. One possible indication that both questions may be researchable is the fact that the low IQ blacks consistently had fewer syllables in mazes than the high IQ blacks or whites. This might mean that the speech of the low IQ blacks is more fluent. One possible explanation, if this should be found to be the case, might be that the language of the low IQ blacks is such that less verbal planning has to be used, i.e., the same structures and patterns are used over and over, with little variety. The question must be researched further.

Two more possible trends are evidenced in Fig. 5. One is that black and white males and females may show different maze behavior, i.e., the white females consistently have more syllables in mazes than the white males, and the black females consistently have fewer syllables in mazes than the black males. This should be investigated more deeply. The second possible trend is that both groups of white retardates appear to decrease the number of syllables in mazes over time while all black groups remain fairly constant. This should be checked out more carefully, and could lead to more interesting results.

Figure 6 presents the data on the percentage of unresolved mazes. Only one comparison is significant: that at T_3 on the IQ breakdown. This is obviously due to the difference between white and black high IQ groups. All groups were able to resolve more mazes at T_3 than they were at T_1. The reversal at T_2 for male and female blacks and whites is interesting, especially in relation to Fig. 5, i.e., white females and black males had

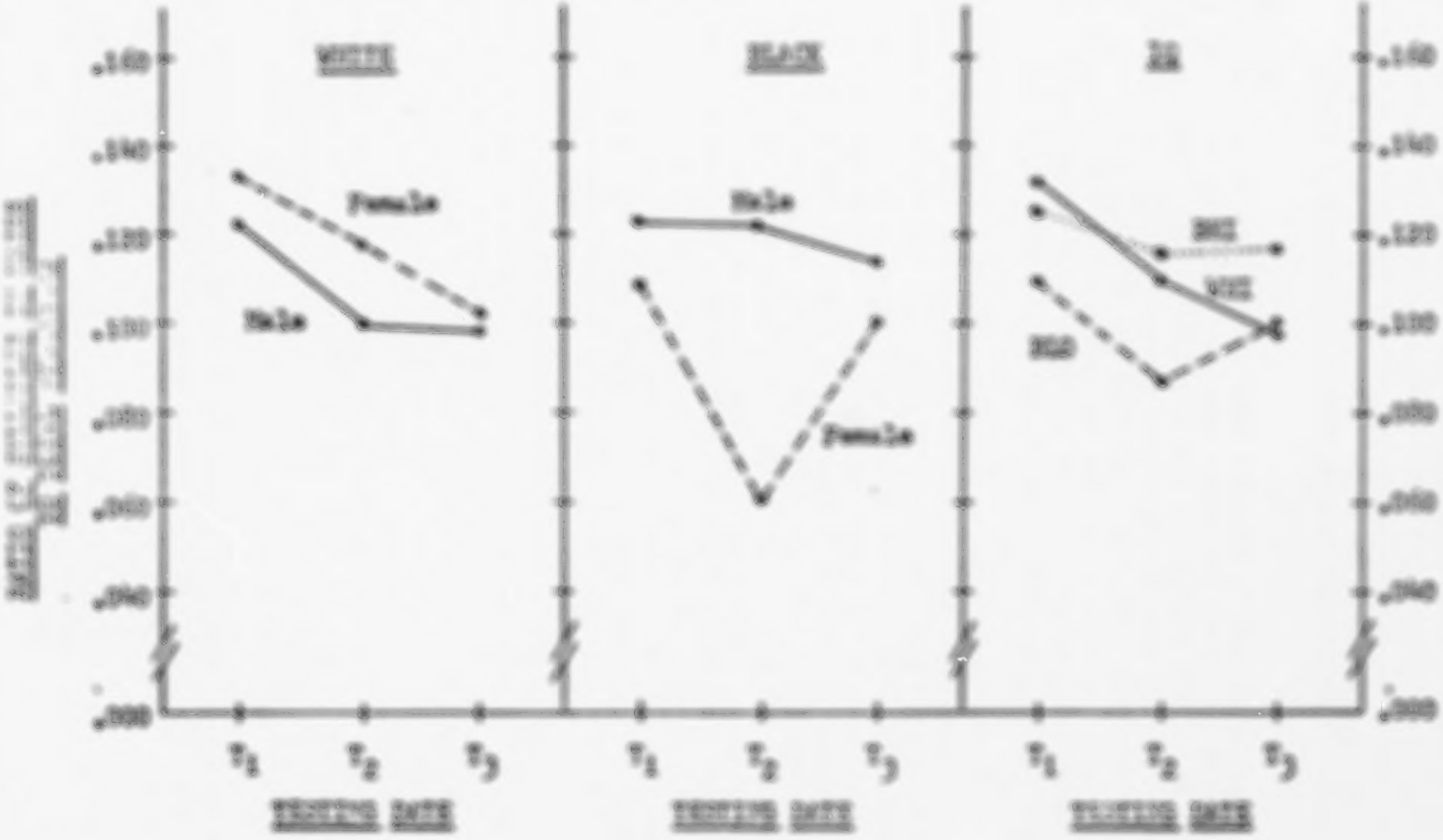


Fig. 3 — Ratio of syllables in mazes to total syllables in transcript as a function of sex, race, and IQ level for [illegible] children over three testing dates.

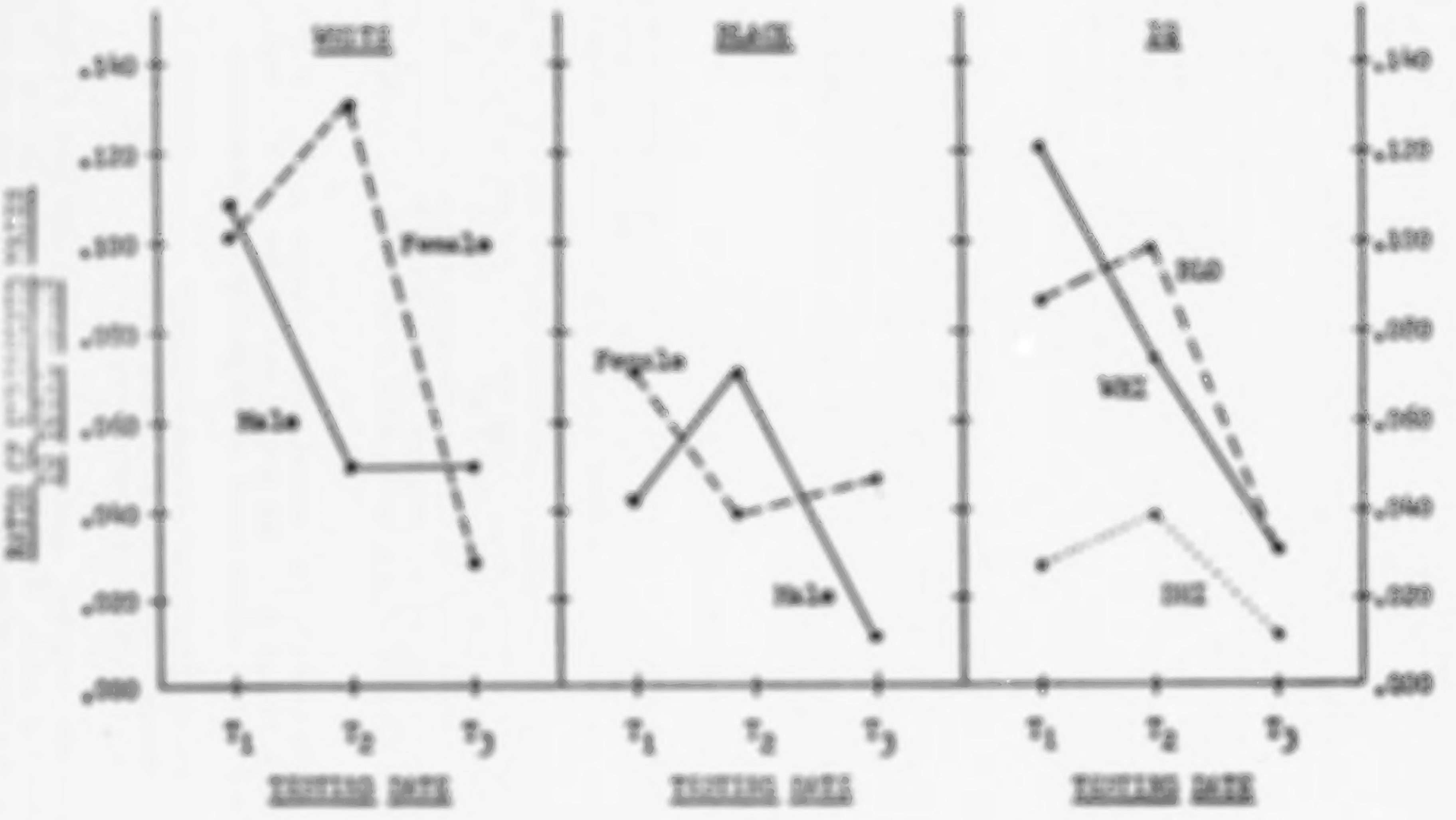


Fig. 6 — Ratio of unresolved cases to total cases as a function of sex, race, and IQ level for EMR children over three testing dates.

were reliable in names, and at T₃ in fact, 3 white females and black males received fewer names. Why this should take place at the time of the first is unclear, but similar sex X race interactions appear at several places in the data and may have more relationship to different role functions of black and white, males and females in linguistic functioning. The fact that the low IQ black retardates were able to produce fewer names than the high IQ black retardates should not come as a surprise. That is, the lower the IQ the more we should expect to see confusion in the search of children. It is surprising, however, to find that the low IQ blacks and the high IQ whites are more similar than the high IQ blacks and whites. Because there are no significant differences at T_1 and T_2 speculation is very hazardous, but it is possible that this might be another indication of the cultural/linguistic factor in retardation. That is, Jensen's term "Environmental Deprivation" may eventually turn out to be more useful than is presently believed.

In the present study, then, the data on names has not been shown to be extremely useful. However, because Loban (1963) had so much success in distinguishing between high and low language-proficiency groups with name data and because there was some indication that high and low retardation might interact with race and sex to some degree, the name measures should be followed up. Other groups of Ss should be added, especially other social class, IQ and racial groupings.

Communication Units, Intonation, and Closure

The communication unit is a linguistic index which was first developed by Loban (1963). The present study, however, utilizes communication units in a way which does not make them comparable to Loban's data. Breaking the single index down into three different indices actually can contribute a great deal to the state of knowledge. Unfortunately, by treating the role of names in relation to communication units, and by expressing the communication units as a ratio of total words rather than as the number of words per communication unit, a great deal of comparative data was lost. However, the measures so gathered do present an original use of the communication unit, and could allow more interpretation to be made which could not otherwise be made.

Figs. 7, 8, and 9 present the data on communication units. It is obvious that there is almost no difference at all in the use of Types 1, 2, or 3 Communication Units by either black or white high IQ retardates. The

Insert Figs. 7, 8, & 9 about here

low IQ black retardates, however, appear to use Type 1 and Type 3 Communication Units to a greater extent than either of the other two groups. This is an indication that the high IQ retardates are less prone to the use of 3-3-1 intonation patterns with grammatically incomplete closure. This Communication Unit is especially prevalent in "after-thoughts," i.e., a statement is made

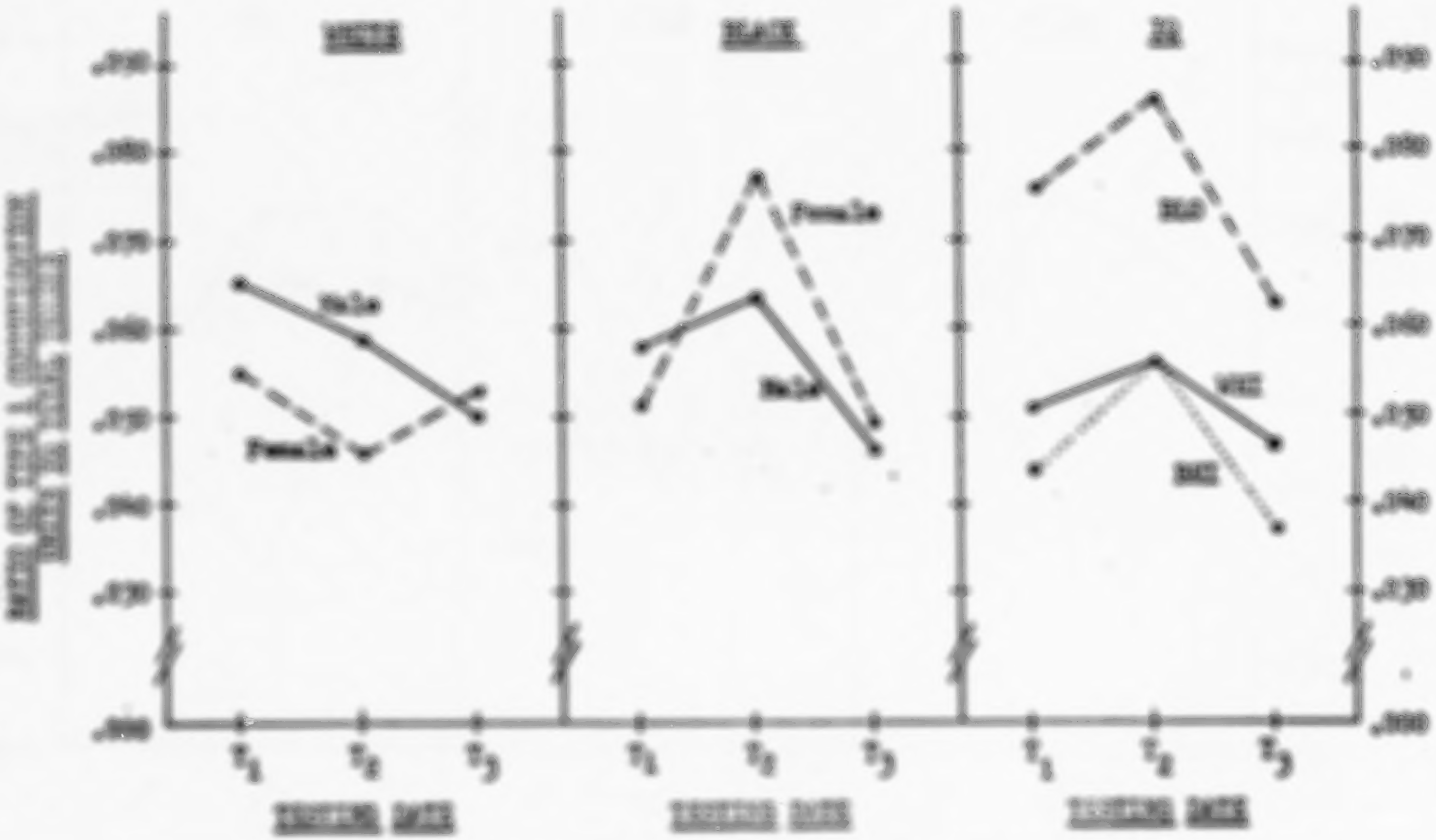


Fig. 7 — Ratio of type 1 communication units to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

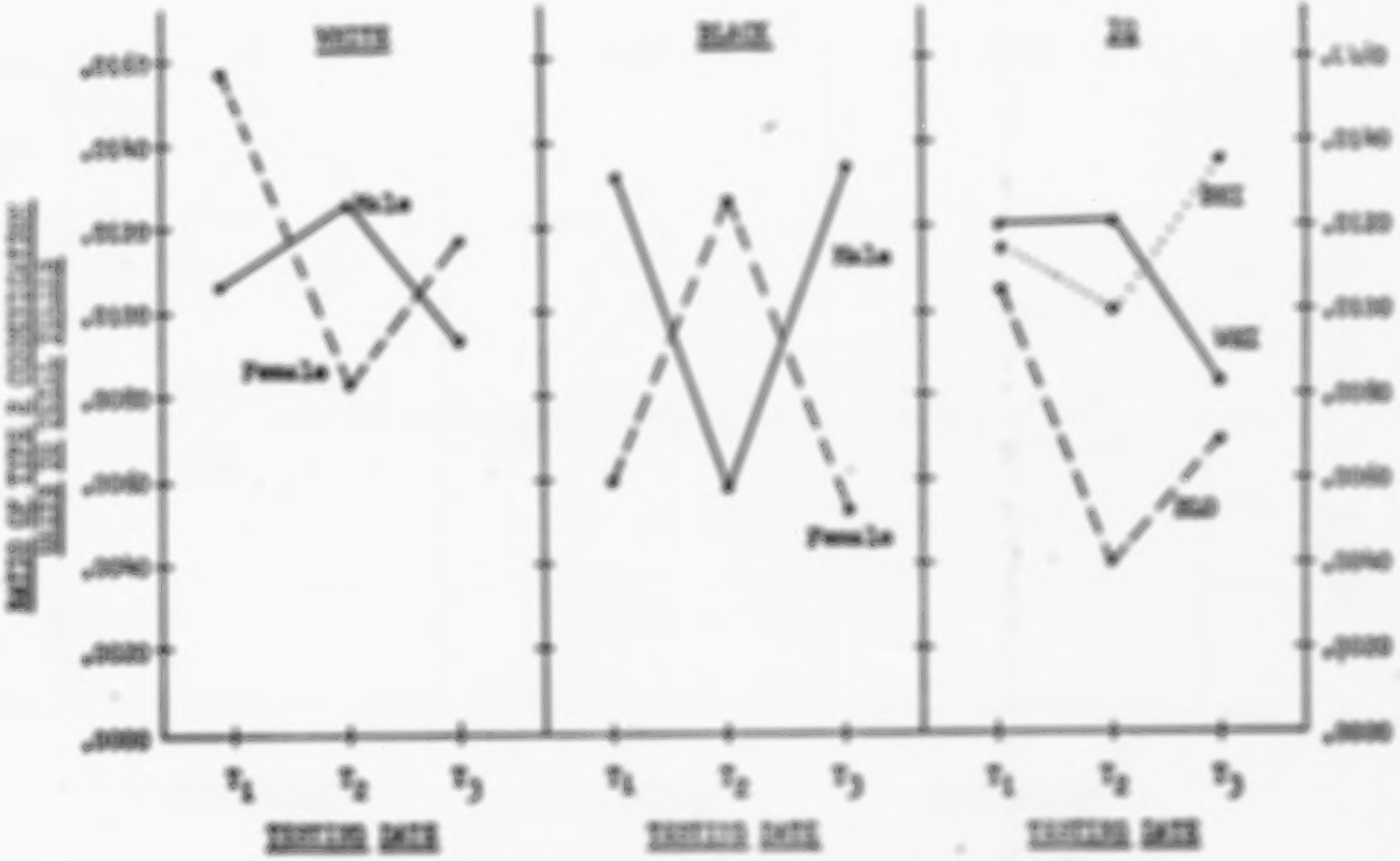


Fig. 8 — Ratio of type 2 communication units to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

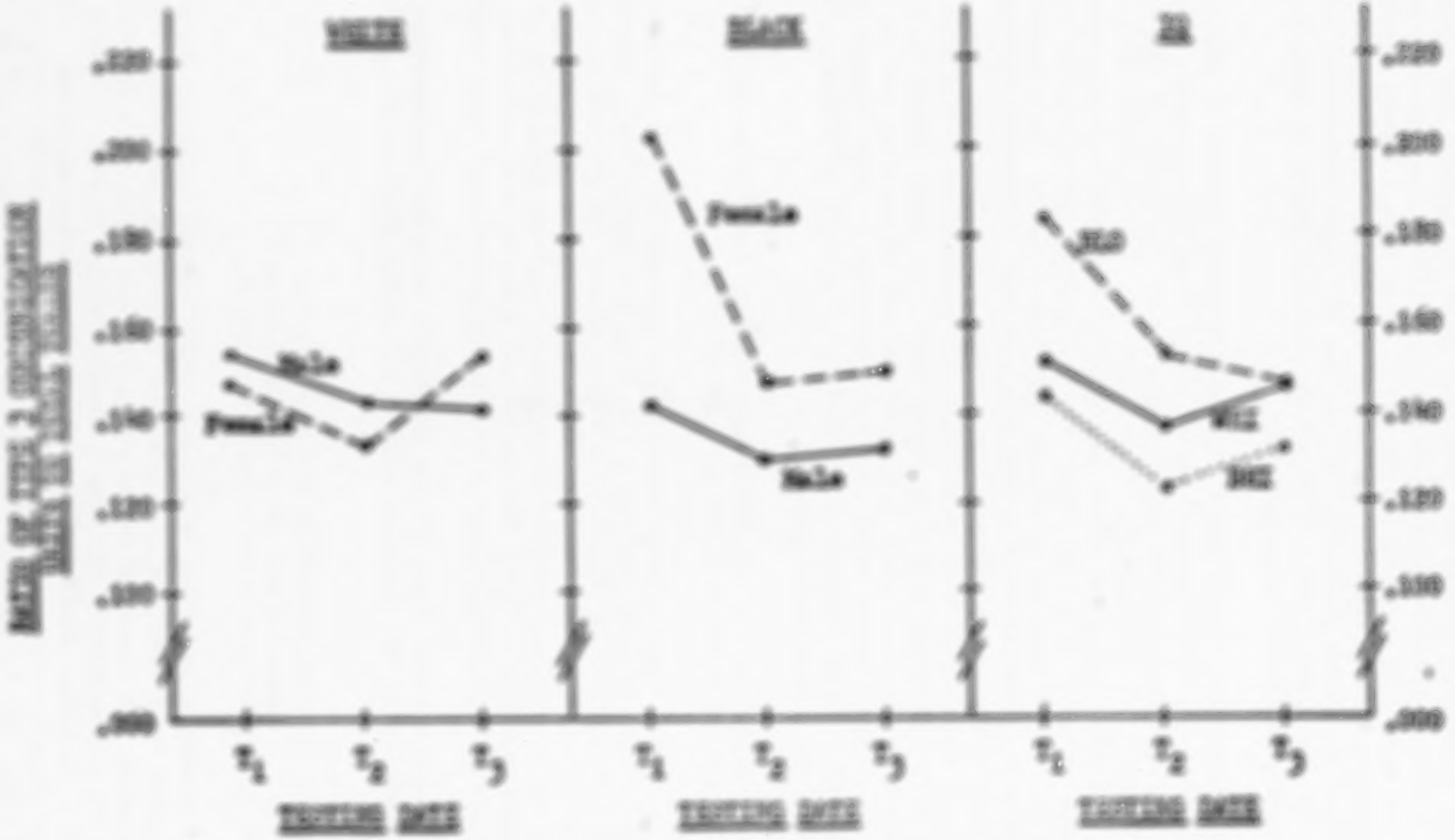


Fig. 9 — Ratio of type communication units to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

with a 2-3-1 contour, and then the same statement is continued as part of the original statement except for the fact that the person had already closed off the first with his intonation pattern. Thus, two utterances appear where one would be expected: the first is a Type 3 Communication Unit. This interpretation could also explain the higher incidence of Type 3 Communication Units exhibited by the low IQ black retardates in Fig. 9. That is, for many Type 1 Communication Units there is also a Type 3 Communication Unit to accompany it.

This interpretation, then, would characterize the speech of the low IQ black retardates as more hesitant, with fewer complex sentences. Note, however, that this interpretation was impossible on the basis of the data presented in Section I. It appears, then, that the low IQ black MR's spoke just as quickly, but that there is probably a characteristic prosodic pattern to their speech. That is, more 2-3-1 contours are probably heard, which makes this speech seem to be choppy and less fluent. This would be only the impression given,however, since the measures of verbal fluency indicate no differences exist between the groups. The fact that both groups of high IQ retardates had fewer Type 3 Communication Units also supports this interpretation since is spent about the same amount of time in the experimental session. Only a certain amount of talking can be done in any given time, and longer, more complex sentences in that amount of time would mean that fewer sentences could be spoken, and therefore, fewer Type 3 Communication Units would appear.

Since there appear to be few differences among the groups in Type 2 Communication Units, nothing will be said about them.

Figs. 7, 8, and 9 also present the racial and sex data on Communication Units. Once more, the graphs suggest a Sex X Race interaction. The lack of significant differences in these measures prevents interpretation without a great deal of qualification. But the fact that a reversal takes place in this set as well as in so many other measures must be indicative of a differential language functioning for males and females from the black and white communities. This possibility should be explored much more deeply.

The data presented in Figs. 10 and 11 should help in interpreting the communication unit data, but unfortunately, it serves neither to confirm nor deny it.

Insert Figs. 10 & 11 about here

Figure 10 presents the data on appropriately spoken complete intonation patterns expressed as a ratio of all complete intonation patterns. The differences between the obtained score and a perfect score of 1.00 would be due, in part, to the existence of Type 1 Communication Units; it would also be due, in part, to names spoken with complete intonation patterns. The two are inseparable in the data as presented. There are no significant differences between the groups, and interpretation of the data is, thereby, severely limited. Once more, however, it should be noted that both black and white

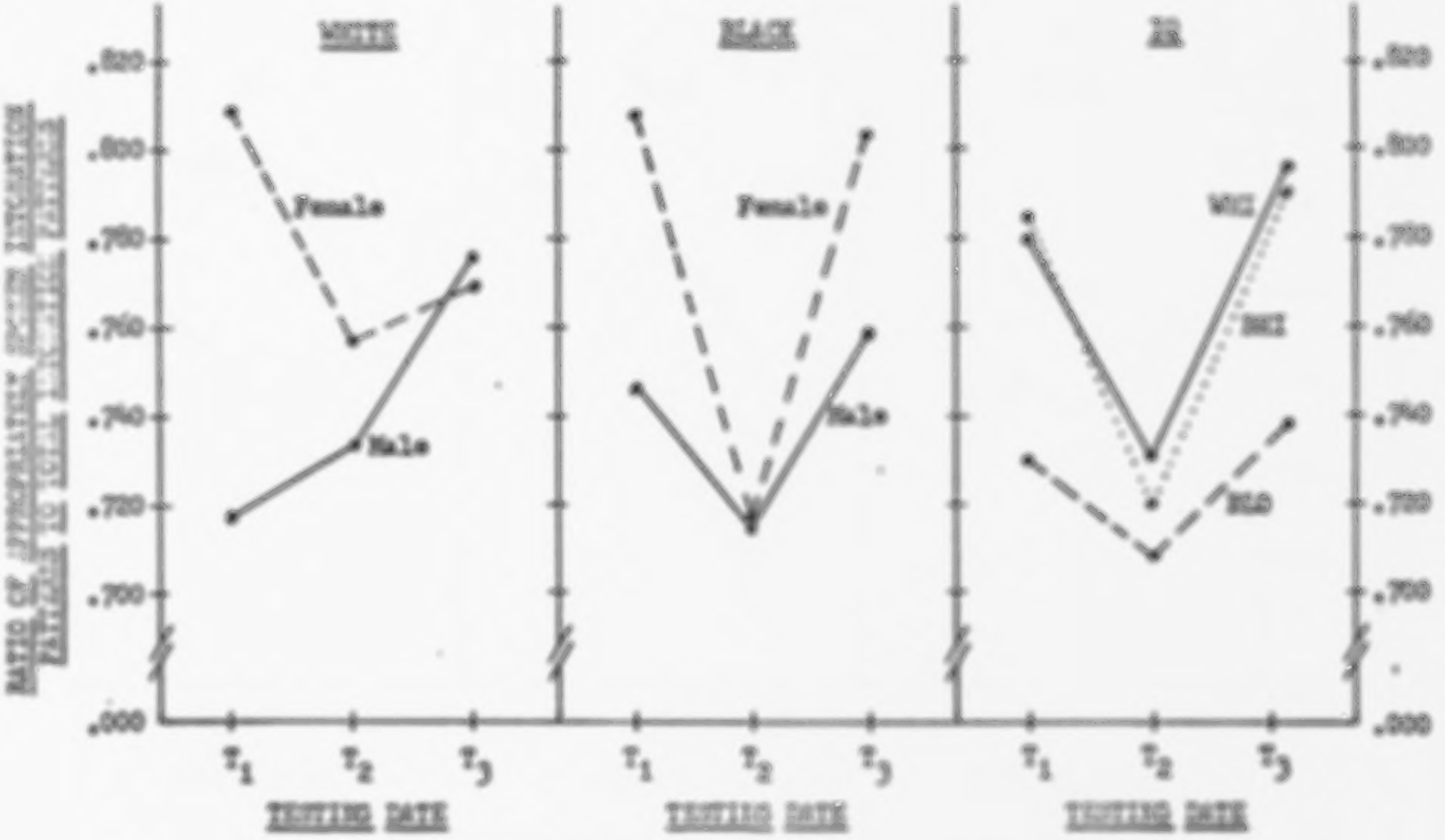


Fig. 10 — Ratio of appropriately spoken complete intonation patterns to total complete intonation patterns as a function of sex, race, and IQ level for EMR children over three testing dates.

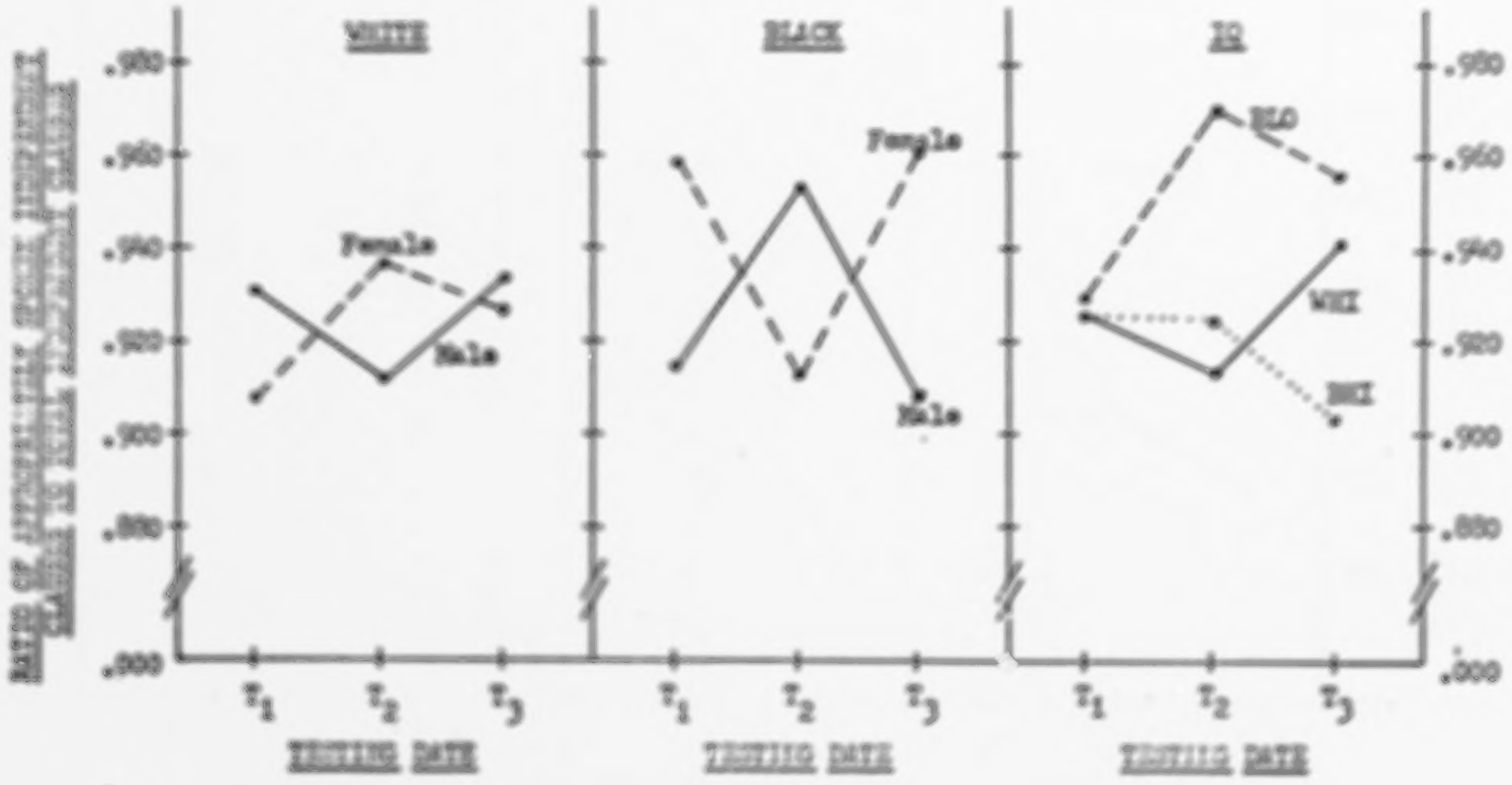


Fig. 11 — Ratio of appropriately spoken independent clauses to total independent clauses as a function of sex, race, and IQ level for EMR children over three testing dates.

high IQ retardates are more similar to each other than are the low IQ black retardates to the high IQ black retardates. This difference, though parallel to the data in Figs. 7 through 9, is extremely small and is simply mentioned.

Fig. 11 presents the data on appropriately spoken complete independent clauses expressed as a ratio of all complete independent clauses. Again, no differences can be said to exist between groups, and all must be said to have been equally proficient in their use of independent clauses. The connection to Type 3 Communication Units should be rather obvious, however, and the slightly higher ratio for black low IQ retardates should be noted as well as the reversal at T_2 for black and white males and females.

Evidently the last two measures are confounded by several factors and their usefulness appears to be severely limited. Either or both might prove useful in contrasting MR's and normals, and SES, racial, and sex differences might appear as well. But for the present subject groupings they appear to have had little usefulness.

In summary, the notion of Communication Units appears to be very useful, and should be studied in more detail. The other two indices have not yet been shown to be useful, although they might well prove to be useful in other contexts.

Dialect Features

The next eight figures can all be thought of as presenting dialect feature data. It is possible for this same data to be seen as representing "errors" in SAE, but since successful arguments have been made for the existence of Black English (Fasold, in press), they will be considered as dialectal differences here.

Black English can be considered to be a different language, or to be a separate dialect of American English. Both arguments have been presented (Loflin & Sabin, 1968; Fasold, 1968); which label is applied is unimportant here. The only relevant fact is that Black English is different syntactically (Wolfram & Fasold, in press) and probably lexically (at least during the early years of a child's life: cf. Entwisle, mimeo). Since this is the case, black children can hardly be scored for "errors" in SAE; rather, the obtained measures in the present study must be presented as differences. It has been pointed out many times that such differences are likely to be markers of black speech (or lower class white speech, as the case may be, i.e., there is no data to suggest that lower class white children either do or do not speak a different dialect from SAE although a hypothesis to this effect is likely to obtain support). It is also true that such markers are likely to work against the social integration of a speaker of Black English (or, if such a thing can be found, lower class White English). As such, the markers must be identified so that SAE features can be taught to the child. Therefore, there is a definite utility in the obtaining of measures such as those found in Figs. 12-18.

There are two major subsections to the dialect features figures. The first subsection contains the acceptability measures (Figs. 12-16), and the second subsection (Figs. 17-19) contains measures on verbs, including "BE." Some of the figures in the first subsection also have elements of the data in the second subsection, i.e., "acceptability" includes notions on verbs as well as on other parts of speech.

Acceptability Measures

Any measure of "acceptability" requires a standard against which to judge it. It is possible that middle class whites and middle class blacks would judge "acceptability" differently in some situations, and this factor needs to be borne in mind. It is also possible for middle class whites to judge the same feature in different ways. For example, the writer's own (Boston) dialect "you hadn't ought to do that" is perfectly acceptable, while to a Midwesterner the same sentence is usually felt to be quite unacceptable. In spite of these cautions, there is likely to be quite a good degree of agreement between middle class raters on the unacceptable features of lower class speech, even if the backgrounds of the raters are markedly different. The Shuy, Wolfram, and Riley (1967) Detroit Dialect Study shows just how accurate such judgments can be. This is likely to be especially true of college-age persons who have grown up in the midst of the nation-wide communication explosion of the past 25 years.

Figure 12 presents the data on acceptable tokens expressed as a ratio

Insert Fig. 12 about here

of all tokens. This ratio includes all of the data contained in the next four figures, and some of the data presented in the following three figures. The data for all groups are very close. The only significant difference is obtained at T_1 on the ID data. It can be seen from the graph that the overwhelming portion of all tokens used by all groups are acceptable, i.e., only a very small percentage of all tokens are unacceptable. This only serves to support the notion that the markers of lower class speech are very few in number, at least on the token level. These few markers, however, are all likely to elicit the "class" response in the middle-class hearer. Only a few features have been shown to be used by persons of all classes to identify speakers from the lower classes (Shuy, Wolfram & Riley, 1967; Tucker and Lambert, 1967). One of the problems in the present study is the lack of studies to which the present data can be compared. It would be very interesting to have comparable data from other cities since the question has been raised to the present author as to whether Black English is as universal as it is generally considered.

Figure 13 presents part of the data subsumed in Figure 11. This data

Insert Fig. 13 about here

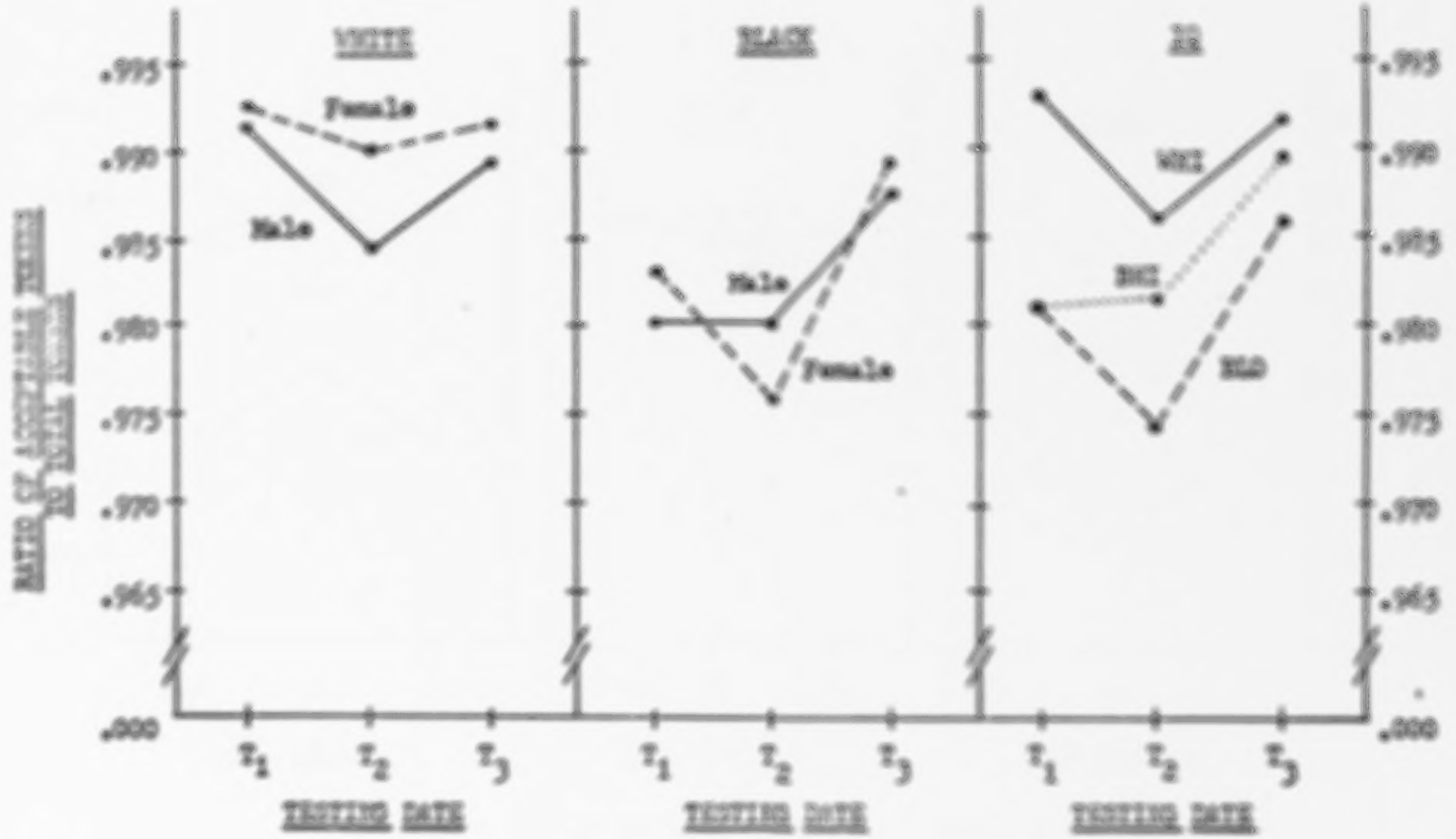


Fig. 12— Ratio of acceptable tokens to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

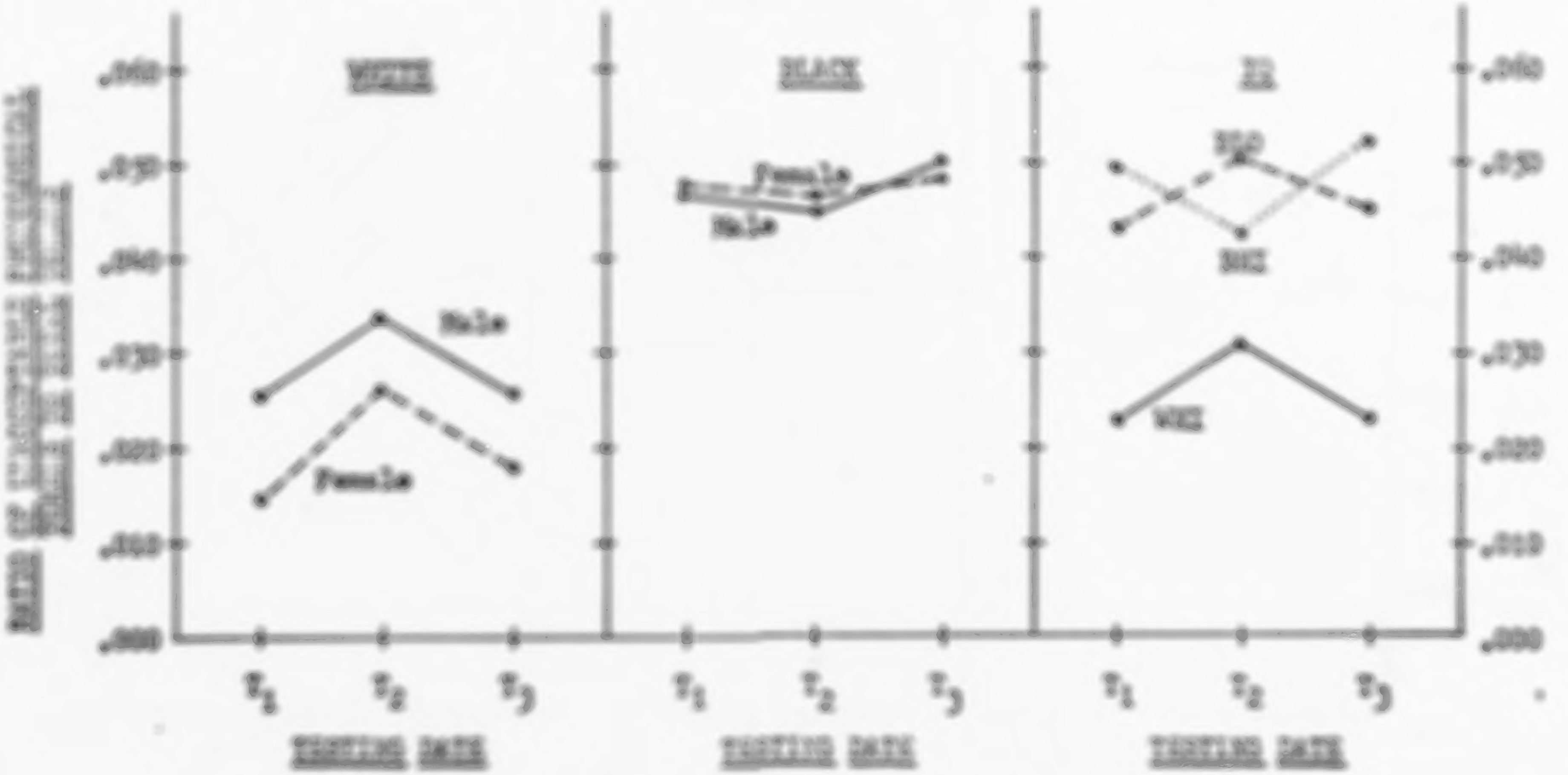


Fig. 13— Ratio of unacceptable phonological tokens to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

is that of the unacceptable phonological tokens expressed as a ratio of total tokens to the total of unacceptable tokens. Greater or lesser differences might have appeared between groups if more data had been gathered, or if the method of gathering data had been different (e.g., recording conversations between the children). As it is, however, blacks, both male and female, and high and low IQ, appear to have more unacceptable phonological tokens than whites in general. Enough work has been done in Black English to corroborate this interpretation. For example, Labov (1967) shows how certain phonological simplification rules are operative in the speec of New York City blacks which require *pass*, *passed*, and *past* to sound identical in Black English. Similarly, *miss*, *missed*, and *mist* are homophonous in Black English. Other phonological rules operative in Black English would produce other "unacceptable" phonological tokens. Some of these unacceptabilities would also have to be taken into account under "unacceptable morphological tokens" (below) because of the interaction with grammatical features apparent in the examples given above. However, these differences are more apparent than real (Labov, 1967) because they are due simply to phonological simplification rules that are operative in Black English rather than to any morphological differences.

It is interesting to note, in connection with Fig. 13, that at no time were the white retardates free from unacceptabilities, although at no time did they have as many as did the blacks. It would be interesting to add middle class white and black comparison groups to this data. The results might give us a clearer picture of the position of retardates in relation to other children.

Figure 14 presents the data on unacceptable morphological tokens.

Insert Fig. 14 about here

Here, again, the differences were very slight; in fact they were even less than those obtained in the phonological tokens measure. It can be seen in the graphs, however, that black children did have a tendency to produce more unacceptable morphological tokens than did white children. Partly, this would be due to phonological simplification rules which operate in connection with inflected morphemes, and partly to certain overgeneralization of inflectional rules, e.g., *mans* for *men*.

Not only were the differences between groups on morphological unacceptabilities few in number, the number relative to the total tokens was quite small. Again, this demonstrates how only a few features out of the total function as class markers.

Figure 15 presents the data on the unacceptable usage tokens expressed

Insert Fig. 15 about here

as a ratio of the total number of tokens. Virtually no differences are apparent between the various groups. Evidently, black and white retardates

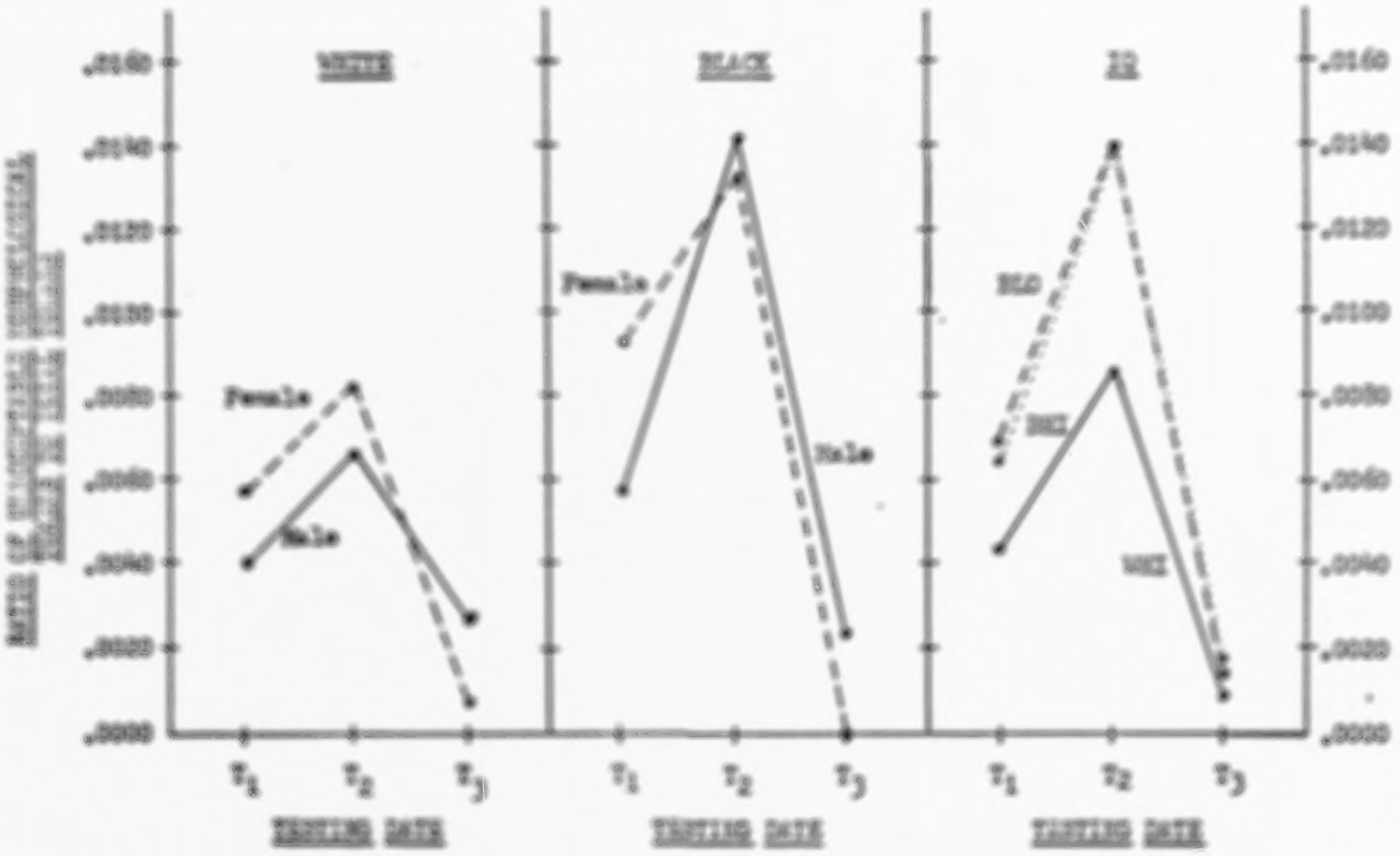


Fig. 14.-- Ratio of unacceptable morphological tokens to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

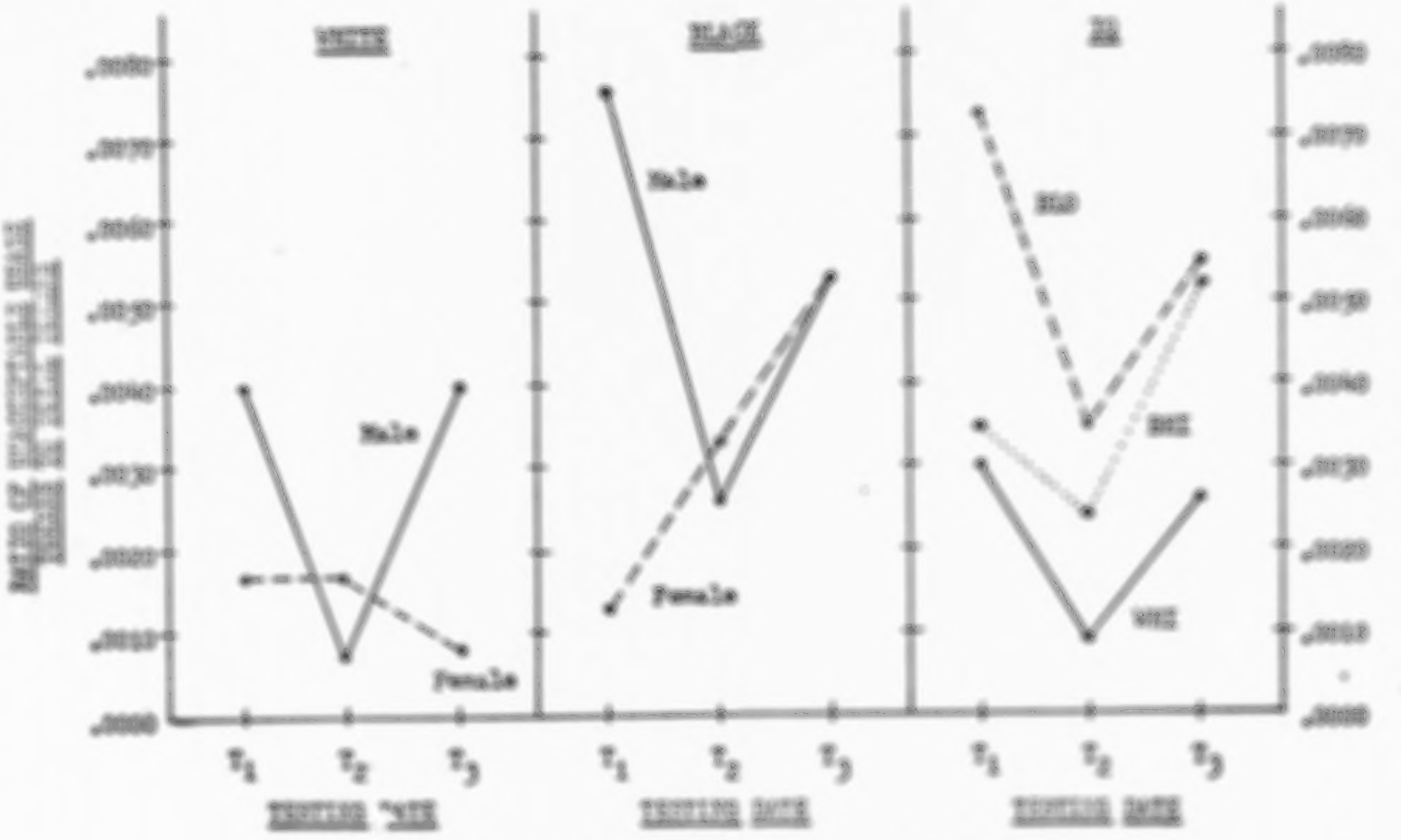


Fig.13 — Ratio of unacceptable usage tokens to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

function at about the same level on this measure. For this data to be meaningful it would have to be compared to equivalent data on normal or middle class children (or both). As it is, the ratios are so small that they preclude any possibility of meaningful comparisons.

Figure 16 presents the data on unacceptable English word tokens expressed as a ratio of total tokens. Again, the expressed ratios are very small, and meaningful comparisons are precluded. Close inspection of the graphs, however, will allow the inference that other comparisons might yield significant results.

Insert Fig. 16 about here

In summary, then, only slight white-black differences appear in the unacceptability data. There appear to be no IQ differences worthy of note, and all of the above differences could be taken as representing dialect features.

Verb Forms

The second subsection is presented in Figures 17-19. One of the greatest areas of differences between Black English and SAE is the verb system, especially as it concerns "BE." That is, Black English (the Washington, D.C., variety, at any rate; there is no guarantee that the same features appear in Milwaukee Black English) has a verb system which contains contrasts such as "He be home" vs. "He home" (Fasold, in press). The first means that he is at home intermittently, while the second means that he is home right now. Therefore, the data on "incorrect" verb forms is quite pertinent, although the use of "incorrect" must be taken with great caution.

Figure 17 presents the data on all incorrect verb forms expressed as a ratio of total finite verbs. This measure contains at least part of the data presented in Figs. 18 and 19. Almost all of the group comparisons are significant except for those between males and females. This is definitely due to the verbal system dialect differences existing between blacks and whites. What these differences might consist of, however, is not available to us from these data because the measure is so gross.

Insert Fig. 17 about here

It is interesting to note that even the white retardates used incorrect verbs as much as .20 of the time (males at T_1). This ratio is surprisingly high, and is indicative of probable dialect features operating in the speech of the whites. This should receive some attention in other studies. It is a very important indication that something is happening. For example, what

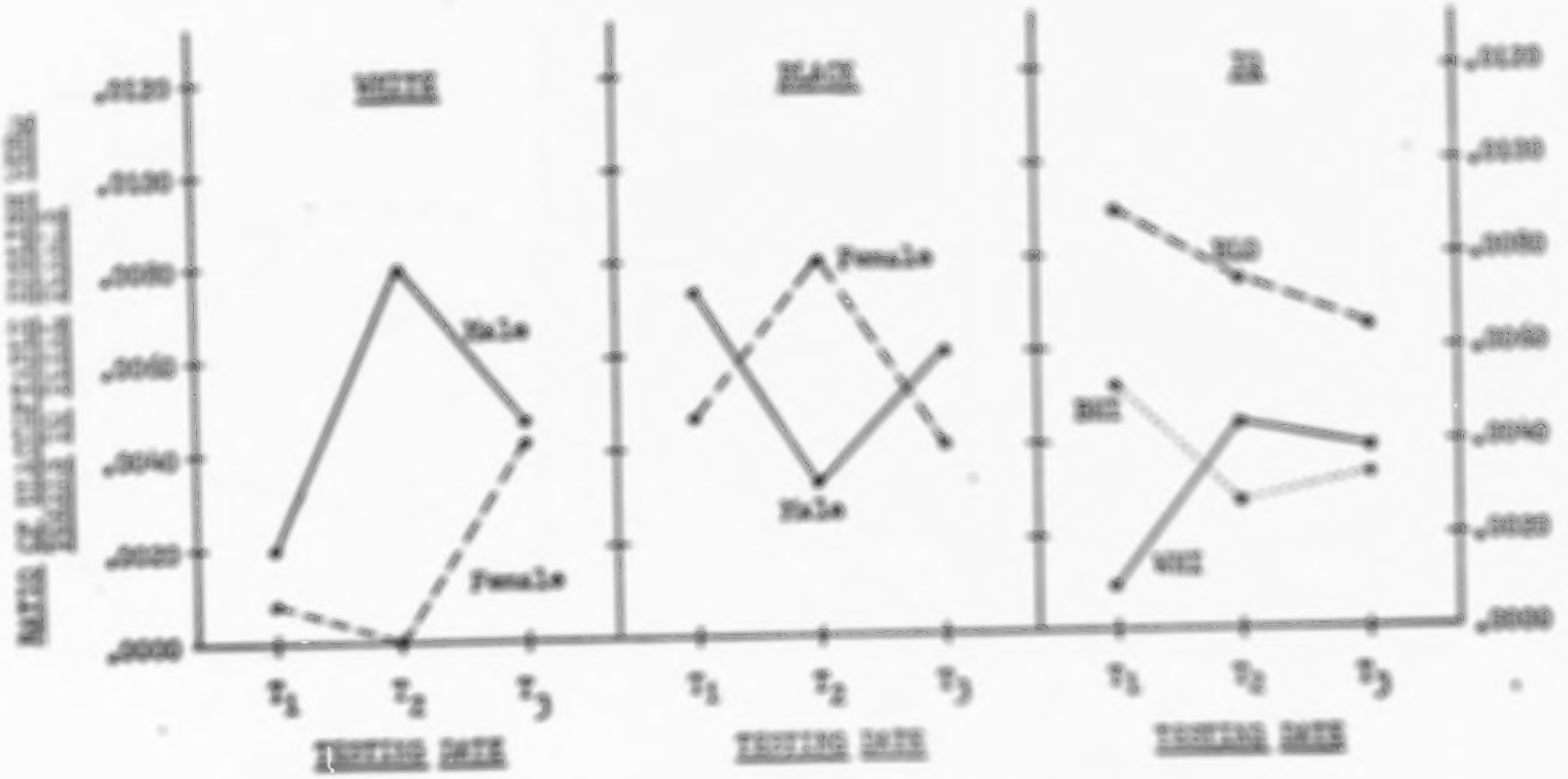


Fig. 16.— Ratio of unacceptable English word tokens to total tokens as a function of sex, race, and IQ level for EMR children over three testing dates.

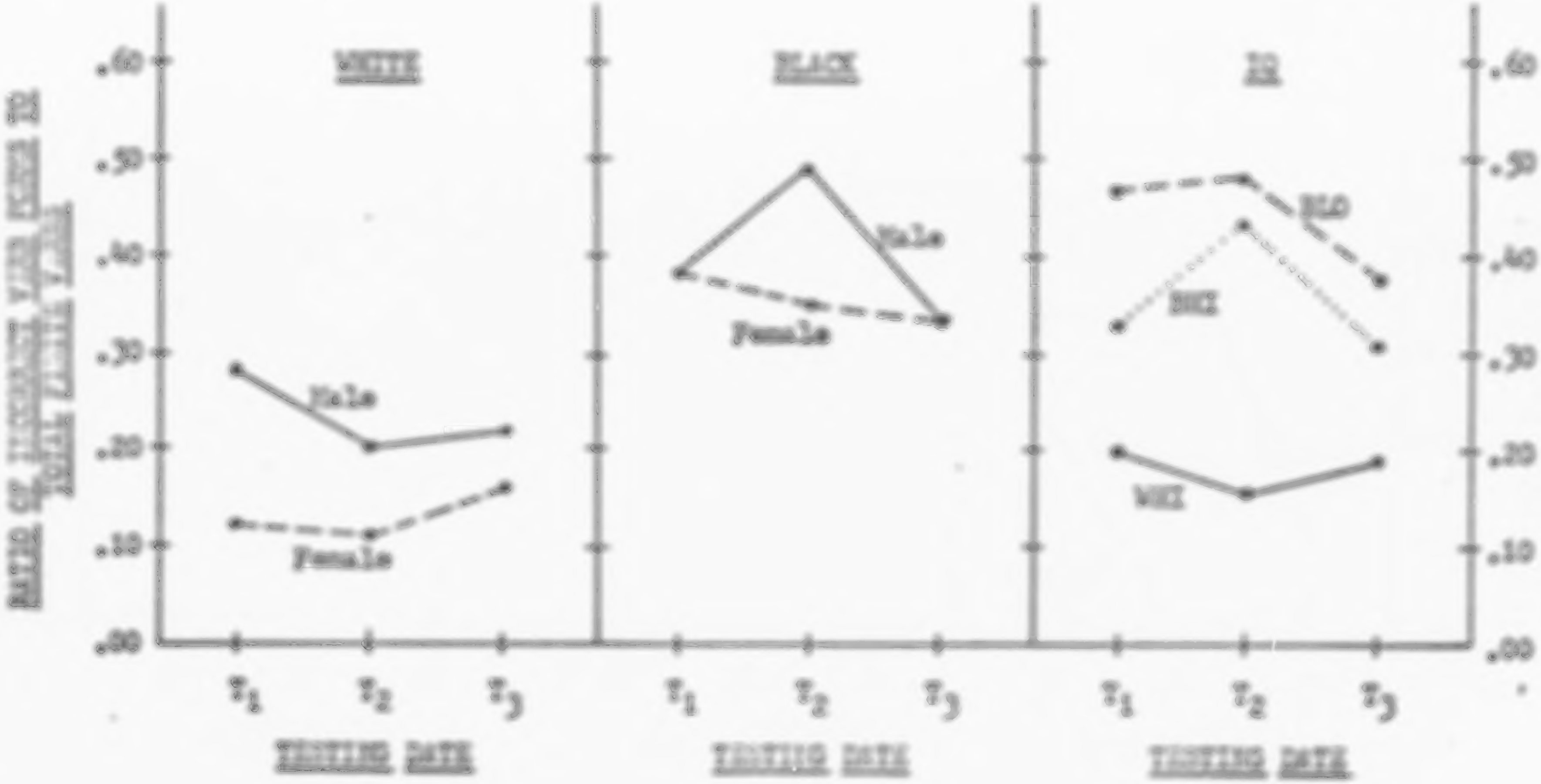


Fig. 17 — Ratio of incorrect verb forms (including omission of "to be") to total finite verbs as a function of sex, race, and IQ level for EMR children over three testing dates.

would be found with other groups included in an IQ X Race X SES breakdown? Would dialect features be associated with MR or SES, or even with both? Speculatively, it would be important to know whether dialect features can be predicted on the basis of a demographic variable such as IQ.

Figure 18 presents the data on omissions of "BE" with the main verb

----

Insert Fig. 18 about here

----

following expressed as a ratio of total finite verbs. Clearly, the blacks do this more often than the whites, although the whites do exhibit this feature at something less than 10% of the time. It is interesting to note that the high and low IQ black retardates do not coincide. This is also true of the male and female data. It is possible that, by CA 11-12, the high IQ black MR's are beginning to use "BE" preceeding the main verb as it is used in SAE. The data are not clear on this, however. The male-female differences seen in Fig. 17 are consistant with Shuy, Wolfram, and Riley's (1967) data on lower-class speech markers, i.e., lower class females tend to use fewer markers than do males. Again, however, the data are not really clear on this point.

Figure 18 presents the data on omissions of "BE" as a main verb. Again, the blacks appear to do this more than do the whites, but it may be that this is due mainly to the effects of the low IQ black retardates. The high IQ black retardates appear to be very similar to the high IQ white retardates in Fig. 18.

It is interesting to note the drop-off for low IQ black retardates at T_3 in Fig. 19. Since language training was one of the instructional goals of this study, it is possible that such training had an effect on this particular dialect feature. If it did, it appears to be the only measure out of the 19 variables which exhibits such change.

In summary, then, these last three figures demonstrate that there are real differences between blacks and whites in their verb systems, but they give little indication of what these differences might be. It is tempting to speculate that they are the same differences as appear in the speech of blacks in New York City or Washington, D.C. This would, however, require assumptions which one might not be willing to make. While the blacks show very strong differences from SAE, the whites also show some differences, and it should not be assumed that the white retardates in this study spoke SAE. There might, in fact, be important dialect differences to be found, and the matter should receive more consideration.

Insert Fig. 19 about here

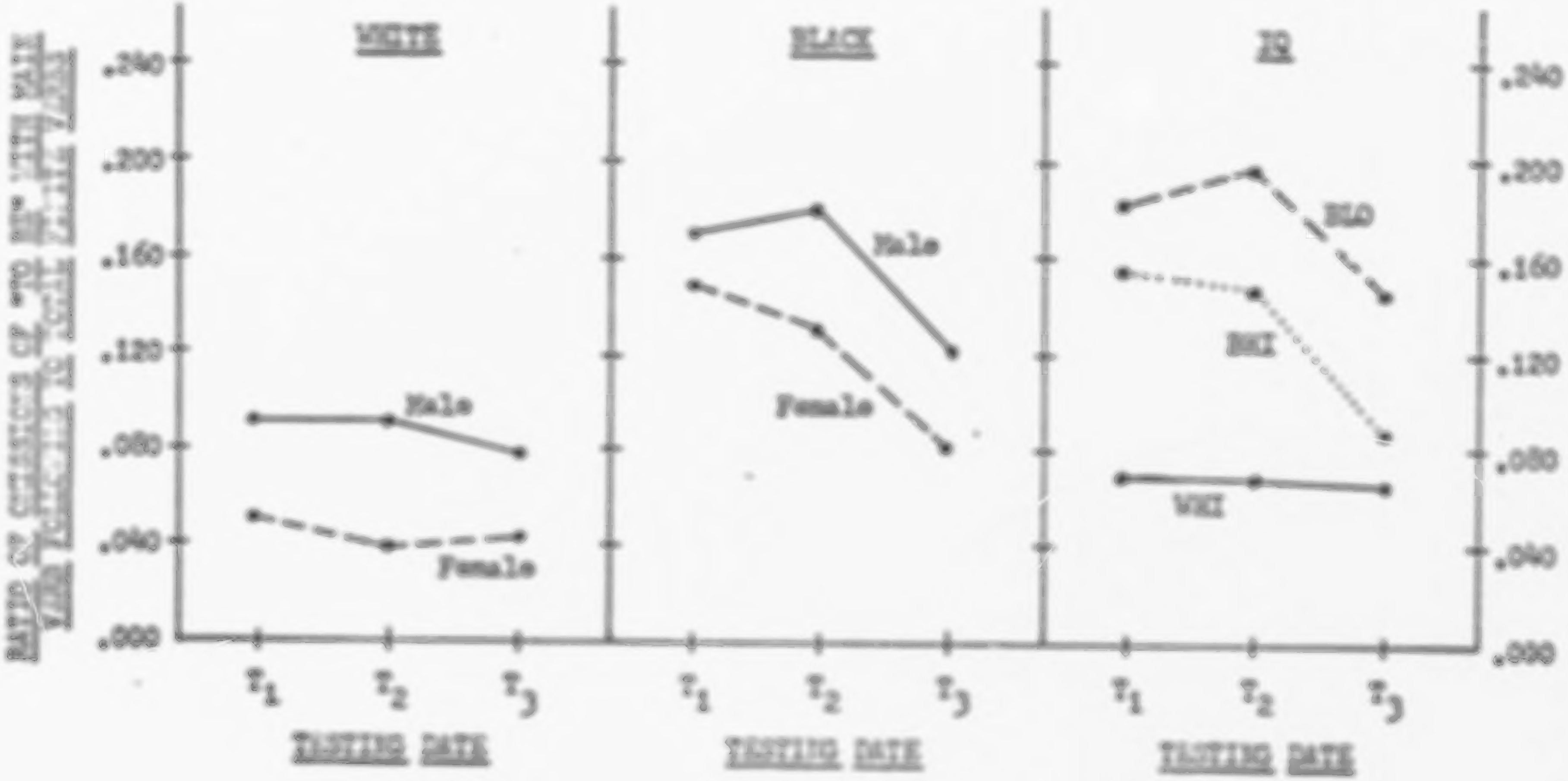


Fig. 18 — Ratio of incorrect verb forms which are omissions of "to be" with main verb following to total finite verbs as a function of sex, race, and IQ level for EMR children over three testing dates.

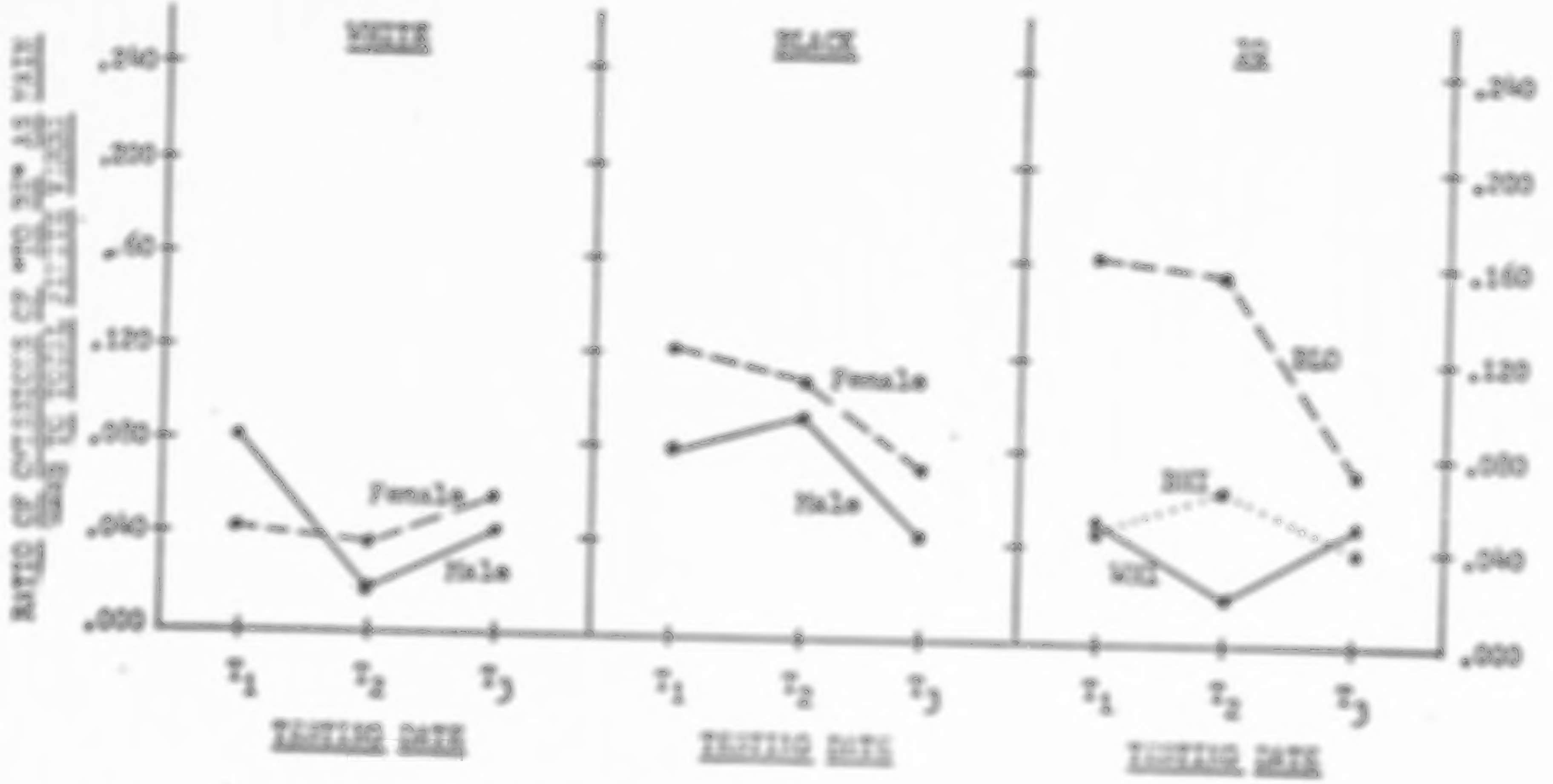


Fig. 19 — Ratio of incorrect verb forms which are omissions of "to be" as main verb (no verb in sentence) to total finite verbs as a function of sex, race, and IQ level for EMR children over three testing dates.

Discussion

The data which has been presented could be interpreted in any one of a number of ways. In another report, Betty Brown considered the measures from the point of view of the clinical speech pathologist. Her report was very much tied to the ITPA, a restriction which severely limited the allowable generality of her remarks. Another writer might consider the same data from the point of view of the behaviorist, and still another might consider them in yet a third way. Each of these viewpoints would cause the resulting presentation to be restricted.

I have tried to present the data within the general framework of socio-linguistics. This also restricted the generality of my remarks. I felt, however, that by presenting them in the manner in which I did, I would be grappling with the basic variables involved in the measures: Race, IQ, sex, and descriptive linguistic measures.

Unfortunately, not a great deal of new information emerged from the analysis. In some cases there appeared to be an hiatus between the present data and the published literature; in other instances new ground was being broken, and in still other instances, old indices were being used in totally new ways.

All in all, the data were remarkably consistent: nothing emerged except for a few random speculations as to what might be obtained under different conditions. An attempt was made to squeeze all of the interpretation possible out of the data, but in general there was precious little to be squeezed.

It appears that some of the measures taken could indeed be very useful. If this is so, why, then, would they fail to be useful in the present study when they were taken three times over an eighteen-month period on the same children? A partial answer to this might be obtainable in the nature of the experimental session itself.

There were three experimental sessions. The first, T_1, took place when the mean age of the children was about CA 11. Most of the $\underline{E}$s collecting the actual data were middle-aged, middle-class, white ladies. This was bound to have an effect on the results because the children were all lower-class children, and about half of them were black. The elicitation of data has been giving socio-linguists the horrors for many years now. It is only recently that methods have been devised which would allow the collection of adequate data. These methods were not used in the present study, unfortunately, and the data is thereby restricted and constrained.

Many things have to be taken into account in the collection of data. For example, one aspect of the black culture is that children are rarely spoken to by adults. They quickly learn to be submissive and evasive in the presence of adults, and to give the appearance of being non-verbal. But when sophisticated methods of data-elicitation are used, it becomes apparent that these children are not only verbal, but that they also have a rich verbal culture (Labov, 1967). Either the same or the opposite situation might exist among lower class white children.

The possibility of a Race X Sex of E crossover interaction was observed by the directors of the project, and male white and black experimenters were added to the eliciting staff at T_2. It is also true that T_2 coincided with the time of the Milwaukee riots. There could well have been a number of Sex X Race X Experimenter interactions, but this was not checked out. It should be. The third session, T_3, included all of the elements of T_2 except for the fact that there were no riots. This would allow a Sex X Race X Experimenter X Emotional Conditions interaction. Still another factor which should be considered in all of this is the fact that the children received 18 months of teaching in this period. There is little indication that the data was affected by this teaching, but this may be due to either the measures obtained or the manner in which they were obtained rather than to any lack of effect. Of course, it is quite possible that there was little or no learning that took place in this period which could affect any language measures at all, let alone the ones which were obtained. This possibility cannot be discounted.

The choice of Ss may have been too restricted. That is, only retardates were chosen for this study, and the delineation into high and low IQ, male and female, black and white may not have been meaningful. What results might have been obtained had normal, middle class, black and white, males and females been added? What if an urban-suburban-rural sample had been obtained? The list could be extended endlessly.

The choice of stimuli for eliciting linguistic samples may have restricted the data unduly. That is, Charles Addams' cartoons were used. These cartoons appear in the New Yorker magazine, and are appreciated by relatively sophisticated adult readers. Could they possibly have not been meaningful to retarded children? A spot-check of the tapes makes one come away with the impression that the children have as much involvement in the action of the pictures as they would in reading a laundry-list. What would have been obtained if a series of Walt Disney cartoons had been used, or a simple face-to-face conversation about last night's TV program (or the race riots)? The Charles Addams cartoons restricted the subject matter to such a degree that certain dialect features could not have been obtained. For example, if a check on verb past tense vs. verb present tense occurrence were run on the null hypothesis that no significant differences would be obtained, it is possible that the hypothesis could be rejected overwhelmingly in favor of the present tense. Lexical dialect features would almost certainly not appear, e.g., "Cool it, Honky."

It is likely that the grouping of children of CA 6-14 might have affected the results obtained. If it is true, as so many writers claim (e.g., Lenneberg, 1967), that grammatical development is complete by CA 6 or so, then perhaps a further breakdown of the data by age is indicated. It is also true that linguistic indices such as the ones used in the present study are likely to be affected by cultural background, the present affectational situation, the present physical situation, etc. They might not be affected by language level per se. This, too, could severely restrict any difference which might have been obtained.

At any rate, the data as presented are not satisfying. I would hope that other work would be done based on the lines of some of the suggestions which I , have offered. If anything, this report raises far more questions than it answers, and to drop the investigation in its tracks would be unfair, both to the present experiment and to the children who were being investigated.

CONCLUSION

Any attempt at generalization, either in terms of subject groups or of groups of indices, is difficult in view of the differing effect of time of these indices across the primary and intermediate subject groups. We may, however, note (1) those indices on which improvement occurred, so that educational endeavor might concentrate on these measures, and (2) those indices on which little or no improvement occurred over time, so that further studies might discover why educational endeavor failed in these areas, a discovery which might lead in turn to improved teaching (or learning) techniques. Measures on which intermediate scores are higher throughout can be considered amenable to change over longer periods of time, for intermediate scores start at a higher base and continue to grow at a faster rate than primary scores. Finally, those indices adversely affected by the riot situation may be regarded as adversely affected by stress or threat situations, a phenomenon of interest to the psychologist.

Improvement was evident across time for both groups of subjects on seven indices, the type-token ratio, number of words per minute, and the proportion of acceptable tokens increased, whereas the incorrect use of to be, both alone and with a following main verb, the ratio of incorrect verb forms to the total number of finite verbs, and the proportion of number one communication units declined. For the primary subjects, we find an increase in the use of appropriate independent clauses and intonation patterns, together with a decline in the proportion of unacceptable phonological and English word tokens. For the intermediate subjects we find a decrease in the ratio of syllables in mazes and of unresolved mazes and in the ratio of unacceptable usage and morphological tokens. The intermediate subjects' scores on unacceptable phonological and English word tokens did not improve, however, nor did their facility in incorporating appropriate independent clauses or intonation patterns. The primary students showed no improvement in the proportion of syllables in mazes nor in number two communication units, though the latter index may be variously interpreted. Neither group improved in the average number of syllables per token or of words per narrative segment, nor was there an increase in the use of type three communication unit.

Whereas vocabulary diversity as measured by the TTR increases with time for both groups, the average number of syllables per token does not enjoy a similar increase, although speech rate as indicated by the number of words per minute also increases for both groups with time, the number of words per narrative segment, an indicator of ideational fluency, does not. Similarly, although the use of intonation pattern alone decreased, we find no increase in the type three communication unit, the apt pairing of 2-3-1 intonation contour with independent clause. Verbal morphology does improve, both and specifically in the incidence of *to be*. The overall improvement in the use of acceptable tokens is reflected in the primary improvement as measured by morphological and usage tokens. The ratios indicating use of appropriate independent clause and intonation pattern increase for the primary subjects alone; the proportion of syllables in mazes decreased only for the intermediate subjects. Nor do the primary subjects decrease their relative usage of number two communications units.

When we turn to intermediate scores as compared with primary scores, we find that on all indices which were improved with educational endeavor for both groups of subjects (TTR, words per minute, number one communication unit, acceptable token, incorrect use of *to be* -- alone or with following main verb, and incorrect verb usage) the intermediate subjects obtained better scores across time. Higher scores were also obtained on the number of words per narrative segment. In these areas, then, improvements may be found not only across time for each group, but by comparing age groups at each period of time. The proportion of unresolved mazes, however, was higher for the intermediate group at each time period, and the number of syllables in mazes higher except at time three. As mentioned previously, the intermediate subjects may be attempting new constructions which they have not mastered by time three. Similarly, the lower scores on number three communication units may indicate use of a diversity of intonation patterns.

Turning to those indices which vary at time two, we find that for both groups the use of appropriate intonation patterns and the correspondent use of number one communication units decreases. The proportion of words per narrative segment declines as well. The proportion of unacceptable tokens, specifically the proportion of morphologically inaccurate tokens, increases. For the primary subjects the use of appropriate independent clauses and of number three communication units decreases, whereas the proportion of unresolved mazes decreases. For the intermediate subjects the proportion of incorrect verbs, specifically that of *to be*, increases, as well as the proportion of phonological and English word errors. The use of number two communication units, those which are grammatically independent but which lack the 2-3-1 intonation pattern, decreases. Under these conditions, however, the use of independent clauses actually increases for the intermediate subjects, as does the type-token ratio. The average number of syllables per token increases for both groups of students. The primary students at time two improve in the proportion of syllables in mazes and in the proportion of unacceptable usage, phonological, and English word tokens.

Such results are quite mixed. The number of words per narrative segment, however, does drop for both groups, an expected result if we assume that the person in a stress situation jumps from topic to topic. The improvement in hesitation phenomena, as indicated by the number of syllables in mazes and the ratio of unresolved mazes, for the primary group is totally unexpected. Similarly, the increase in syllables per token for both groups and the increase in the type-token ratio for the intermediate group at this time, supposedly indicators of vocabulary fluency and diversity, were also unexpected. Further studies may determine whether such findings hold true only for this group of subjects or for others matched as to age, sex, race, and degree of mental retardation.

Portions of a Dialogue Between Two Linguists

That the field of linguistics is in a state of flux, or perhaps better stated another way, that the indoctrination process in linguistic departments is not complete, can be evidenced by the following exchanges of comments. Dr. Dever's paper, just preceding, was reacted to by Dr. Kitonga. Her reactions were then commented upon by Dever, whose comments in turn elicited further comment by Kitonga. At the time of writing, it would seem that the female of the species (Kitonga) has had the last word.

The exchanges are herewith presented with no comment.

I. Kitonga Reaction to Dever Paper

Section #1

p. 1 - Clarify what Bernstein is getting at. The distinction between *what* is said vs. *how* it is said seems to indicate a difference in suprasegmentals or extralinguistic features such as "tone of voice," facial expressions, etc.

Expand on "more fluent and structurally predictable." How is this tapped by objective indices? (lexical fluency by TTR, structural by subordination index, disfluency by maze indices?)

Rate wpm - Rapid speech may here be equated with fewer hesitation phenomena, but generally people speak more rapidly to "make up" hesitation phenomena. (Goldman-Eisler, 1956, Agnello, 1963) The difference here may be due to coding of mazes, filled hesitation phenomena, but not pauses; the latter were found to be the prime determinants of rate.

The difference in rate may be due to the fact that whites may be learning the conventional response, which they give automatically, whereas the blacks have a greater discrepancy between "home" and "school" speech, and have not yet learned the automatic response in the latter environment.

p. 2 - Why the "loosening-up effect" for blacks alone? The black rate, again, may be continuing to grow because they have not yet learned, or go beyond, the expected and predictable verbal response. The sex difference is a sensible aspect of societal roles, for it is the black females and white males who are expected to be aggressive, hold the family together, etc.

words per narrative segment - If we can view this as a rough index of organization by topic, then it assumes great importance. Williams and Naremore (1968) note (p. 19) that "In many of the papers cited earlier in this report, one of the most consistent social class differences in modes of speech was the degree of organization which united the segments of a message."

p. 3 - Perhaps "taciturnity" is the wrong word. It seems that this index reflects the tendency to speak incoherently, i.e., jump from topic to topic when excited. If, however, as Goldman-Eisler (1961) claims, unfilled pauses reflect a cognitive form of speech, filled pauses an emotional form, it is strange that the mean rate in Table 3 shows a higher incidence for black males and white females.

syllables per token - For this sample at least, there is a significant correlation with TTR, another measure of the same aspect of speech. We may choose whichever is easier to code.

p. 4 - This analysis of the use of Chas. Addams cartoons assumes that the subjects respond to this stimulus as we do. We must recall that the picture is not presented as a cartoon. If bizarre situations are depicted in these pictures, stranger situations are encountered daily in TV programs and comics. The difference would seem to be cross-cultural instead. Get some pictures of ghetto life to respond to if you want involvement.

p. 5 - Clarify the paragraph explanation of comparisons with Labov's BE grammars.

syllables in mazes

p. 5 - Blacks may have more mazes because they are attempting more complicated structures beyond the simple semi-automatic and stereotyped response.

unresolved mazes - There may be more syllables in mazes and fewer resolutions of these mazes because of passivity and corresponding frustration in white females and black males, but this speculation is very tenuous.

p. 6 - As for Environmental Depression, get a black cartoonist.

INTERVIEW

This whole section might be organized around "fluency" vs. "verbal output." Words per minute might be a gauge of hesitancy vs. rapidity of response, words per narrative segment a gauge of "ideational fluency" (Carroll, 1964), syllables per token and TTR a measure of lexical fluency, and maze indices measures of disfluency.

The idea of black language/dialect might be examined more closely, in socio-psychological terms as well as linguistic. Grier and Cobbs (1968) suggest that in the black adult at least, use of the dialect may serve (1) to reinforce stereotypes and thereby soften hostility or (2) as a demand acceptance. The psychological part of this study may suggest whether both, or either of these devices are being used by the subjects. The idea of sex role and the function of language therein also demands further attention. Are other traits, such as aggressiveness, measured in the psychological subsection?

II. Communication Units, Intonation, and Clauses

p. 6 – "unfortunately" may be replaced by "however."

Define "after-thoughts" more closely. To me at least the phrase implies a higher level of cognitive functioning and organization than found here.

Aren't type one communication units added to type three? You seem to imply that the situation is the other way around by your phrasing. These categories may be roughly equivalent to those of Williams and Naremore (1969), who obtained similar results with a cross-status comparison. Although their results are recorded in terms of percentages (page 22), equivalent figures seem obtainable here.

p. 7 – You note that the use of 2-3-1 contours "makes this speech seem to be choppy and less fluent." Is this true if they are after-thoughts, as you suggest? Perhaps content is a compensatory factor.

p. 8 – The reversal at T_2 may again reflect sex role. Speech may be more disconnected in times of stress.

III. Dialect Features

p. 8 – What's the Fasold reference? Sounds interesting.

The issue of black (or lower class white) speech raises several competence-performance questions. Does this speech reflect differences in linguistic competence or performance alone? If the former, at what level are the grammars different? If the latter, is the performance lacking a result of a paucity of situations which elicit such speech? (Williams and Naremore, 1969, pp. 60-61).

Acceptability

p. 9 – If the clang response is at work, counting the occurrence vs. nonoccurrence of features may be acceptable and save coding time, as in my study.

p. 10 – "Greater differences might have appeared between groups if more data had been gathered." Why assume this; less differences might also have appeared.

It is strange that morphological differences should prove less important than phonological. Francis (1955) noted that social differences were reflected mainly in morphology and syntax, and Pederson (1964) noted that verb morphology was most important in distinguishing standard from nonstandard usage. Putnam (1955) and Williamson (1964) mentioned the importance of verb morphology, especially to be, in describing Negro speech as you note on the next page. The distinction between linguistic, statistical, and psychological significance applies here.

Verb forms - Stewart (1964) points out the difference in to be + VP as aspectual.

p.11 - "What these differences might consist of is not available to us from these data because the measure is so gross." See my breakdown.

p.12 - The effect of formal training on dialect is negligible, at least in my study, where different kids are tested at the same time vs. the same kid across time. Both groups show an increase with age in the occurrence of to be + VP, subject omission, and substitution of base for past. This increase is especially evident in Negro speech. Errors in the use of to be in the past tense quadruples with age in the black population. The $[T_2]$ inversion is erased among non-blacks by intermediate age, whereas black usage more than doubles in moving from primary to intermediate. Similarly, the white use of the double negative decreases with age, whereas the black usage more than doubles.

The use of here/there + plural is a white feature. In fact, each primary student had the feature whenever there was a chance of its occurrence.

Discussion

p.14 - Again, try cartoons by a black cartoonist. The TV program discussion idea sounds best. Williams and Naremore (1969) found it better than games or aspirations as a topic, both in terms of the amount of speech produced and in distinguishing between groups by age, sex, and race. The race riots topic seems too dangerous. You note the invalidity of data gathered at T_1. It might follow a cartoon or television discussion so that different forms of speech are elicited. (Labov, 1966, DARE)

The linguistic environment, or chance of occurrence is indeed important. I have used the same in my study.

ANALYSIS

Williams and Naremore (1969, p. 4) note the problems in dealing with "a great variety of detailed linguistic and production characteristics" is that such studies "become unwieldy inventories of detailed characteristics, where if the balance of items tipped one way or another, a label of the code was assigned. Much of the criticism of the empirical research into codes of speech rests on this problem." They cite Lawton's Social Class Language and Education (N.Y., 1968). If this is one of your Lawton references, bring out this point (unless, of course, you interpret Lawton differently).

II. Comments on [illegible] Reaction to [illegible] Paper

1. P.17 - "The difference in rate may be due to . . ." But there was no difference in rate. In spite of the fact that the graphs show differences, there were no statistical differences, and anything that appears on the graph is a function of the scale used.

   Because both groups were lower class, and because both groups were retarded, and because there are no systematic significant differences between the groups, it is hopeless to try to speculate on the "differences between groups." These "differences" are not "due to" anything because they do not exist.

2. P.17 (mention of p. 2) - I cover the social role aspect.

3. P.17 (VPVS) - The data are still incomplete and I can't comment. It may be that this is quite important, but you can't say so without the data to support it. See my comment - final sentence, next-to-last paragraph on page 3 of my manuscript.

4. P.18 - (mention of p. 3) O.K., I'll go along with this, but how do we know that jumping from topic-to-topic is a way of filling pauses? It could well be that it is a tight-lipped refusal to speak. Just because the mean data for black males and white females is highest does not mean that we can connect anything. Let us not forget that the VPVS data is incomplete.

5. P.18 (syllable-token comment) A significant r between TTR and SPT is fine, but see my comments on TTR. I wonder if the TTR in this study can tell us anything? If not, where does this leave the SPT measures?

7. P.18 (p. 4) - My comments on the use of the C.A. cartoons stand. I have had much better luck with other things. I don't know if pictures of ghetto life would help matters. Is involvement really what is needed, or is the establishment of a flow of conversation more important? There have been a number of good methods of data elicitation discovered, any one of which would have been better than the one used. The final comment doesn't really make sense.

5

8. P. 18 (p.3) - Do you still have the original? Something was left out of the final paragraph here that was in the copy I handed to the typist. I don't even understand what I meant.

9. P. 18 (p.3) - But blacks do not have more syllables in pauses.

10. P. 18 (comment on unresolved pauses) - This speculation is extremely tenuous because of the lack of significant differences.

11. P. 18 (p.4) - This comment is naive.

12. P. 18 (overview) - Great theorizing, but the lack of significant differences makes this difficult to say. I really think I milked this section for all it was worth.

13. P. 19 (p.6) - My statement says that type I are added to type III. Unfortunately I don't have the Williams and Naremore paper. Complete the reference - they have a number of publications together.

14. P. 19 (p.7) - I note that the use of more 2-3-1 contours makes the speech seem choppy and less fluent. This stands to reason: the more sentence contours in a given length of time, the more choppy the speech would seem. Granted, it could also be more fluent in some interpretations, but let us not forget what the tapes sound like in the first place (at least, the ones I got to listen to).

15. P. 19 (p.8) - Again, this could be interpreted like this, but dammit, the differences are non-significant, and any interpretation is not only tenuous, it is dangerous. I think that I may already have overstepped the bounds of propriety with the interpreting I have done, and I don't want to go any further. The most that I would allow myself on these measures is to note the continuously recurring patterns!

16. P. 19 (p.8) - See reference list.

The next set of questions, while interesting, are quite unanswerable. Competence is a notion that must be handled in terms of generative grammars, and the present study was one in terms of structural grammars, and thereby limits the possible interpretations to those of only performance. I have written a working paper on this, if anyone is interested.

17. P. 19 (p.9) - Probably.

20. P. 21 (p.11) - O.K., but my statement reads "from these data." This statement and [illegible]'s statement will make good, complementary reading.

21. P. 20 (p.14) - Isn't it too late?

22. P. 21 (Lawton Ref.) - Different Lawton. See reference list.

III. [illegible] Comments on Dever Defense

The basic point at issue seems to be my temerity in speaking of nonsignificant differences. Dever chooses to limit the meaning of "differences" to "systematic significant differences" -- see, for example, his opening comment. Accepting the sacred cow of statistical significance, Dever takes no note, beyond a passing reference (XIII)[1] to my paper, of linguistic as opposed to statistical significance, although he himself referred to the "'slang' component" in his original paper (p. 9 - acceptability measures). Certainly the importance of binary occurrence - nonoccurrence *versus* quantitative cumulation is a nontrivial question in the consideration of linguistic performance.

This author would, of course, agree with Dever (13) ". . . the differences are nonsignificant, and any interpretation is not only tenuous, it is dangerous. I think that I may already have overstepped the bounds of propriety with the interpreting I have done." My comments were, therefore, carefully phrased as suggestions, not assertions, by the use of "if," "perhaps," "it seems," etc. Speculation, once its limitations are admitted, would seem to be as admissible in the [illegible] reaction as in the original paper.

Beyond the basic issue of linguistic-statistical significance, the other points at issue may be seen as the result of either (1) the practical difficulties or (2) misinterpretation of my comments. Under the first category would fall the typographical errors or omissions (Dever 2) and the inaccessibility of the bibliographies appended to the original paper and the reaction thereto (Dever 13, 16, and 22).[2] Dever's misinterpretation of my comments seems but another manifestation of whatever underlies his interpretation of my comments as unsupported decrees *ex cathedra* as opposed to his own carefully delimited and tentative suggestions. Thus he fails to read my comments in terms of the original paper. To my suggestion that cartoons by a black artist might be used, he replies, "Isn't it too late," failing to note that this comment was included as a suggestion for future research, as were his comments, (page 16 of the original final section, first two complete paragraphs) to which I was reacting.

---

[1] References to Dever's reaction are indicated by the number of the statement therein.

[2] Hopefully the consistent misspelling of "consistent" and its derivatives is due to mechanical error. The Williams and Naremore article is "On the Functional Analysis of Social Class Differences in Modes of Speech," Institute for Research on Poverty: Oct. 1968.

This suggestion seems to be a sore point, for Dever has mentioned it twice before - Dever 7 and 11. Statement seven again reveals Dever's failure to refer back to the original. Thus, he asks "Is involvement really what is needed, or is the establishing of a flow of conversation more important?" "Involvement," however, was his term (original, page 16 of the final section) and my suggestion was prefaced with "If you want involvement." His final comment "The final comment doesn't really make sense" doesn't really make sense to me. The immediate response is not only "Why?" but "Which final comment?" If he is speaking of involvement, the problem is of his own making. There is no contradiction between his discussion and mine, for we both offer criticism of and substitutes for the use of Charles Addams cartoons. If, however, he is commenting on the initial comment, it is explained in the following two sentences.

Dever 11, "This comment is naive.", refers to my As for Environmental Depression, get a black cartoonist." His comment is, in turn, empty, for it fails to indicate what areas are encompassed within the scope of my naivete. If the reference is to my interpretation of "Environmental Depression," I admit myself at a disadvantage, since I did not have a copy of his bibliography. If, however, the point is again that of using a black cartoonist, such artists do exist (see, for example, the Bootsie cartoons of Ollie Harrington), and their work might well be employed in such a study. By incorporating such prompts the environmental aspects of the stimulus would be subject-centered and, hopefully, more productive in terms of linguistic output.

On two points, however, I stand corrected. Dever's 13 refers back to my "Aren't type one communication units added to type three? You seem to imply that the situation is the other way around by your phrasing" and asserts that "My statement says that type I are added to type III." In "Two utterances appear where one would be expected: the first is a type 3 communication unit and the second is a modifying clause which must be counted as a type 1 communication unit", Dever does give type three communication units temporal priority, as they were given conceptual priority, in the previous sentence. What I was specifically objecting to was the seeming contradiction in the final (summary?) sentence of the paragraph in which the precedence of the two types of communication units seems to be reversed: "That is, for many type 1 communication units there is also a type 3 communication unit to accompany it." (My italics.) Interestingly enough, Dever failed to catch substantial errors in my comments on Rate one. Rapidity of speech corresponds directly to the amount of time spent in hesitation phenomena, not the converse, and interphrase pauses were found by Agnello to be the prime determinants of subjective measures of rate, not objective measures thereof.

REFERENCES

Bernstein, B. Elaborated and restricted codes: their social origins and some consequences. Amer. Anthrop., 1964, 66 (6, part 2), 55-69.

Cowan, P. A., Weber, J. A. P., and Klein, J. Mean length of spoken responses as a function of stimulus, experimenter, and subject. Child Develop., Vol. 38, No. 1, 1967, 191-203.

English, H. B. and English, A. C. A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. New York: Longmans, Green and Co., 1958.

Entwisle, Doris R., Semantic systems of children. In Frederick Williams (Ed.), Language and Poverty: Perspectives on a Theme. Chicago: Markham Publishing Company, (pre-publication edition).

Fasold, R. Isn't English the first language, too? Paper presented at the Annual Conference of the National Council of Teachers of English, Milwaukee, Wisconsin, November 29, 1968.

Fasold, R. Tense and the form "be" in Black English. Washington, D.C.: Center for Applied Linguistics, Language (in press).

Labov, W. Some sources of reading problems for Negro speakers of non-standard English. In A. Frazier (Ed.), New Directions in Elementary English. Champaign, Illinois: National Council of Teachers of English, 1967, 140-167.

Lawton, D. Social class differences in language development: a study of some samples of written work. Lang. and Speech, 1963, 6, 120-143.

Lawton, D. Social class differences in group discussion. Lang. and Speech, 1964, 7, 183-204.

Lenneberg, E. Biological Foundations of Language. New York: Wiley & Sons, Inc., 1967.

Levin, Harry and Silverman, Irene. "Hesitation Phenomena in Children's Speech," Language and Speech, VIII, 1965, 67-85.

Loban, W. The Language of Elementary School Children. NCTE Research Reports #1. Champaign, Ill.: NCTE, 1963.

Loban, Walter. Language Ability Grades Seven, Eight, and Nine. Washington, D.C.: U.S. Department of H.E.W., Office of Education, 1966.

Loflin, M., and Guhin, R. Transformational rules and competing generalizations. Paper presented at the Annual Conference of the National Council of Teachers of English, Milwaukee, Wisconsin, November 29, 1968.

Maclay, Howard and Osgood, Charles E. "Hesitation Phenomena in Spontaneous English Speech," Word, 15, 1959, 19-44.

Robinson, W. Elaborated code in working class language. *Lang. and Speech*, 1965, *8*, 243-252.

Shuy, R., Wolfram, W., & Riley, W. Linguistic Correlates of Social Stratification in Detroit Speech. Final Report, CRP #6-1347, USOE, 1967.

Trager, George L. and Smith Henry Lee, Jr. *An Outline of English Structure (Studies in Linguistics*, Occasional Papers, No. 3), Norman, Oklahoma: Battenburg Press, 1951.

Tucker, G. R., & Lambert, W. E. White and Negro listeners' reactions to various American-English dialects. McGill University, 1967 (mimeo).

Wolfram, W., and Fasold, R. A Black English translation of John 3:1-21; with grammatical annotations. *The Bible Translator*, in press.

APPENDIX D

Testing Procedures and Copies of Forms
Used in the Study

1. Analysis of Errors in Using ITPA

2. Modified WRAT Administration Instructions

3. Forms Used in Study

1. *Analysis of Errors in Using ITPA*

During the initial testing in the project, the logistical problems involved in testing some 270 children loomed rather large. By accident it occurred that one pupil was given the ITPA twice on the same day while two others were given the test twice three weeks apart. Apparent score discrepancies between the two occasions resulted in a detailed analysis of the records of the three pupils.

Table 1 present in summary fashion the ITPA raw scores of the three pupils obtained on test and retest. In the table and in the subsequent analysis, the letters A, B, and C refer to the three pupils while the Roman numerals I, II, III, IV, and V refer to the five testers involved. These testers were all school psychologists certified by the state of Wisconsin except tester I, who had administered some 100 ITPA's primarily within a research context.

Insert Table 1 about here

Table 1 shows very marked discrepancies between test-retest scores on both the total score and the overall differences of the subtest scores. Some of the sources of error within the test-retest situations are due to changes within the subject from one time to the next; differences between two testers in terms of biases, experience with the test, and the particular type of tester; ambiguity of test scoring and administration instructions; and situational differences. In the analysis of the score differences, it was assumed that the major source of error was ambiguity of test scoring and administration instructions interacting with certain deficiencies in training and experience of the testers.

Table 1

Display of ITPA raw scores obtained on test-retest of three EMR pupils by five different testers

	Pupil and Birthdate					
	[illegible]		[illegible]		[illegible]	
	Test [illegible]	Retest [illegible]	Test [illegible]	Retest [illegible]	Test [illegible]	Retest [illegible]
Examiner No.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ITPA Subtests: Auditory Vocal Automatic	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Visual Decoding	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Motor Encoding	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Auditory-Vocal Assoc.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Visual-Motor Sequencing	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Vocal Encoding	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Auditory-Vocal Sequencing	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Visual Motor Assoc.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Auditory Decoding	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
ITPA Total	106	99	101	118	137	180
Total score test-retest difference	7		17		43	
Subtest score differences ignoring direction	47		36		47	

Auditory Vocal Automatic Test

Pupil A: 1st Test = II   Retest = I
          Score 3   Score 9

Several times the child gave the correct verb but its incorrect form such as crash for crashed, big for biggest, dress for dresses, hat for hats, and these answers were counted wrong. (Some of this might have been their dialectic dropping off of the last part of these words.)

Pupil B: 1st Test = IV   Retest = V
          Score 14   Score 10

This child, on the same day, used entirely different words several times, to obtain the correct score.

Example:	1st time	2nd time
	crash =✓	crashed = OK
	big =✓	larger = OK
	big one =✓	smallest = OK
	painter man =✓	painter = OK
	steal =✓	stole = OK

One tester gave credit for faulty yet understandable speech (fief/thief), yet the next took it away for spelling (thiefes/thieves).

Pupil C: 1st Test = III   Retest = V
          Score 10   Score 17

One point is given for each correct answer. The first tester did not credit item 10 (painter-man) yet the retester (V) credited "painter." The first tester (III) should have stopped after six incorrect, yet after item 13, seven were incorrect, and then the last two items were given (21 and 22) and these were OK and counted!

Visual Decoding Test

Pupil A: 1st Test = II   Retest = I
          Score 14   Score 10

Here no improvement was made by Pupil A on the retest score. In fact three consecutive items were correct on the first test (items 16, 17, 18) and not even given on retest, because of four consecutive failures (items 13-16).

Some responses were scored differently.

1st tester, item 9 incorrect, 2nd tester, correct

1st tester, item 11 correct, 2nd tester incorrect

1st tester, item 13 correct, 2nd tester incorrect

Pupil B: 1st Test - IV Retest - V
Score 14 Score 10

After item 10, six consecutive answers were incorrect and V stopped; IV stopped after item 18. However, all previous items but two (14 and 17) were correct! We don't know what time of day these two examiners administered these tests, i.e., which tester was first. (Tester IV was arbitrarily chosen as first.)

Pupil C: 1st Test - III Retest - V
Score 14 Score 17

Here the obvious retest improvement on retesting is three points. Neither tester encountered six consecutive mistakes. On the retest, the pupil got items 17 and 26 wrong, although these had been correct before. He corrected himself on items 9, 11, 13, 18, 20, and 24, however.

Motor Encoding Test

The purpose is to determine subject's ability to supply appropriate gestures, no verbal response. The scoring here appears difficult, as parts or (partial) gestures aren't given a maximum score, and thus the examiner must judge them.

Pupil A: 1st test - II Retest - I
Score 8 Score 13

Recount of retest score should have been 14. In the retest, the pupil corrected himself and improved on items 3, 4, 8, 12, 13, 14. The pupil did not retest incorrectly on any of his 8 original correct test gestures.

Pupil B: 1st Test - IV   Retest - V
Score 26   Score 17

It is not known which of these tests was administered first. If V was first, then improvement was evident. In the V test, item 3, 11 and 16 went ungestured, and only one part of these items was gestured: 6, 8, 9, 10, 14, 15.

Pupil C: 1st Test - III   Retest - V
Score 16   Score 19

The difference here was small, yet items 10, 13, and part of 14 were correct on the first test, but not on the retest. Items 3, 5, 7 were gestured correctly on the retest, after no gesture on the first test.

Auditory Vocal Association Test

The purpose of this test is to assess subject's ability to relate verbal symbols - meaningfully, by analogy. Testing stops at six consecutive failures.

Pupil A: 1st Test - II   Retest - I
Score 7   Score 10

Items 3, 6, 16, and 18 were incorrect on first test, but corrected on retest. Yet item 20 was correct at first, but incorrect on retest. The total difference and the above information seem to indicate some small improvement.

Pupil B: On the IV test, score should add up to 19, not 20. Also on this test, items 6, 7, 8, 9, 12, 14, 18, and 22 were correct, while on the V test, these were incorrect. Item 2 was correct for V, yet not for IV.

Pupil C: 1st Test - III   Retest - V
Score 11   Score 21

His retest showed a correction of items 7 and 12, plus those six consecutive incorrect (and 2 additional); therefore, a retest improvement.

Visual Motor Sequencing

Purpose of this test is to assess subjects' ability to reproduce a sequence of visual stimuli from memory (picture and geometric forms used) (and test until three consecutive items are failed). Two points for first trial, one point for second trial.

Pupil A: 1st Test - II Retest - I
Score 10 Score 10

Both test and retest identical in score and items passed!

Pupil B: 1st Test - IV Retest - V
Score 13 Score 13

The manual (p. 44) asks that a / be placed on the appropriate line if item is passed, and a 0 if failed. Neither examiner did this. Both scores are the same, yet not all of the items are. For example:

Tester	Trial 1	Trial 2	Item #
IV	0	0	6
V	0	/	6
IV	0	/	7
V	/	0	7
IV	0	/	8
V	/	0	8
IV	/	0	10
V	0	0	10
IV	0	/	11
V	0	0	11

Pupil C: 1st Test - III Retest - V
Score 8 Score 14

These examiners also did not use the / and 0 as requested. One used P's and the other + and -. Although there was a six point raw score retest improvement, they were not in the same areas. See below:

Tester	Trial 1	Trial 2	Item #
III	0	/	2
V	/	0	2
III	0	/	3
V	/	0	3
III	0	/	4
V	/	0	4
III	0	0	5
V	0	/	5

cont.

Tester	Trial 1	Trial 2	Item #
III	[illegible]	[illegible]	8
V	[illegible]	[illegible]	8
III	[illegible]	[illegible]	11
V	[illegible]	[illegible]	11

For III, 3 consecutive failures were 8, 9, 10

For V, 3 consecutive failures were 11, 12, 13

Vocal Encoding

Test purpose is to determine number of unique, meaningful ways subject can verbally characterize a simple object (ex. ball, block, etc.). Time allowed is one minute, and no gestures acceptable as in motor encoding test. The subjects' questions (Q marked) on the form, have an effect on the score. Responses are either simple or related (these are repeated, redundant, etc., responses). Each scorable response gets a maximum of one point credit. This is not an easy test to score! Since it is so conversational, the degree of rapport with the examiner could influence differences in score.

Scores that were obtained from the three pupils on the two occasions were patterned as follows:

	Item 1 ball	Item 2 chalk	Item 3 block	Item 4 cellophane	Total	
1st test - II	4	2	3	2	11	Pupil A
2nd test - I	3	2	[illegible]	[illegible]	[illegible]	
1st test - IV	6	6	6	3	21	Pupil B
2nd test - V	6	6	[illegible]	2	[illegible]	
1st test - III	6	3	5	4	18	Pupil C
2nd test - V	3	3	6	2	[illegible]	

The discrepancies here are obvious suggesting consistent biases or tolerances on the part of the testers.

Auditory Vocal Sequencing

The purpose of this test is to assess subject's ability to reproduce a sequence of auditory stimuli from memory. Once again, the manual suggests

/ and 0 for failure. Two points for first trial passing, and one point for second trial, stopping after three consecutive failures.

Pupil A: 1st Test = II   Retest = I
Score 13   Score 14

The one-point difference came on item 6. The subject was able to repeat the digits on the 1st trial for two points, but he lost one point on item 8.

Pupil B: 1st Test = IV   Retest = V
Score 17   Score 17

Here scores are the same, but see example:

Tester	Trial 1	Trial 2	Item #
IV	/	0	6
V	0	/	6
IV	/	0	8
V	0	0	8
II	0	/	9
V	/	0	9
IV	0	0	10
V	/	0	10

Pupil C: 1st Test = III   Retest = V
Score 14   Score 20

The retest improvement came twice in the first trial and twice in the second. First tester III had three consecutive failures for items 9, 10, and 11, then retester V picked up two points on 9 and 14 and one point on 10, 13. (15, 16 and 17 were his three consecutive failures).

Visual Motor Association

The purpose is to assess ability to relate visual stimuli meaningfully (pictures and objects are used).

Pupil A: 1st Test = II   Retest = I
Score 9   Score 14

Items 8, 9, 12, and 18 were improved on the retest. I stopped after 19 and 20 were incorrect. II stopped after item 17, but 17, 16, 15, 14, 13, and 12 were incorrect!

Pupil B: 1st Test = IV   Retest = V
          Score 18   Score 13

The manual says (p. 31) if items 5, 6, and 7 are passed (which they were for both testers), to continue testing (beginning with item 8) until the ceiling is reached or test ends. The ceiling is reached when two in any eight consecutive items are passed. However, V recorded only to item 21 (14, 15, 16, 17, 19, and 20 were incorrect, 18 was correct). IV completed entire test, listing all failures of 20-28 (except for item 23). None of the failures were the same for test and retest!

Pupil C: 1st Test = III   Retest = V
          Score 20   Score 23

These items were incorrect on both tests: 21, 22, 26. All 28 items were given. Retesting suggested improvement.

Auditory Decoding

The basal level here is a bit confusing; manual says (p. 33) that if subject passes five or more of first eight items, then to continue to test till ceiling is reached (when four in any eight consecutive items are passed) or test ends.

Pupil A: 1st Test = II   Retest = I
          Score 25   Score 5

On retest, nine items were given, four passed, so test should not have been stopped. Therefore, five is an incorrect retest score.

Pupil B: 1st Test = IV   Retest = V
          Score 24   Score 31

Tester IV administered the entire test, while V stopped after item 30, even though there were many correct responses (for example, items 28, 27, 24, 23, 22, 21). Considering the items of both tests from 20 and up, item 26 was incorrect at both times; and items 19, 20, 25, 29, and 30 were incorrect for V's test.

Pupil C: 1st Test = III   Retest = V
            Score 24   Score 31

V didn't give the last item (36) and III gave all 36. The following are items incorrect on both tests: 28, 32, and 35. Retest seems to be an improvement.

Such gross errors and discrepancies as have been noted caused a reappraisal of the approach to obtaining measures of the dependent variable. A number of strategies evolved. First, a consistent administrative and scoring scheme was developed (see attached paper). Secondly, all ITPA's were rescored using the consistent approach. Third, post-testing was done using testers trained by the project personnel to be consistent and reliable in their administration and scoring.

Basals and Criteria for Stopping the Administration of ITPA Test

1) Auditory-Vocal Autom.

Ceiling - 6 consecutive items failed
Basal - 6 consecutive items passed
Record - subject's word response
Score - 1 pt. for each correct response

note: Test intent is to assess grammatical, not articulatory ability.

2) Visual Decoding Test

Ceiling - 2 correct in any consecutive 8, or end of test (whichever reached 1st)
Basal - 4 consecutive items passed
Record - subjects point to selection; examiner records # of selection on comparison sheet
Score - 1 pt. for each correct # response

3) Motor Encoding Test

Ceiling - None } Administer entire test
Basal - None }
Record - gestures only (not verbal responses)
Score - 1 pt. for each gesture part or gesture type (0 if no gesture offered by subject)

4) Auditory-Vocal Assoc. Test

Ceiling - 6 consecutive items failed, or end of test reached (whichever is first)
Basal - 6 consecutive items passed
Record - subject's word response
Score - 1 pt. for each correct response

5) Visual-Motor Sequencing Test

Ceiling - 3 consecutive items failed on both 1st and 2nd trials, or till end of test is reached
Basal - 3 consecutive items passed on the 1st trial
Record - ✓ when item is passed and a 0 when item is failed
Score - 2 pts. for items passed on 1st trial
1 pt. for items passed on 2nd trial

6) Vocal Encoding Test (difficult to score)

Ceiling - None} Hand child each of the four items separately and
 } let him speak for one minute
Basal - None}
Record - 1. subjects verbal responses (should sound like conversation)
2. ✓ the common characteristics mentioned on record form
3. subjects questions as Q
4. underline subjects reponse to examiners Q
Score - No subject has received more than 27 pts.

7) Auditory-Vocal Sequencing Test

Ceiling - 3 consecutive items failed on both 1st and 2nd trials; or till end of test
Basal - 3 consecutive items passed on 1st trial; or till end of test
Record - Place a check if digit is correctly repeated; 0 if incorrect
Score - 2 pts. for correct response on 1st trial
1 pt. for correct response on 2nd trial
note: Digits for this test should be uttered by examiner at 2/sec.

8) Visual-Motor Association Test (difficult to score)

Ceiling - A. If MA is 6-0 and passes 2 or more object items - give pic. items. Then ceiling is reached when 2 in any 5 consecutive items passed.
B. If MA 6 or over, start with item 5. If subject passes item 5, 6, and 7, continue till ceiling (when 2 in any 8 consecutive items passes or test ends). If subject fails item 5, 6, or 7 continue test till ceiling (when 2 in any 8 consecutive items are passed or test end)
Basal - (give progressively easier items, beginning with 4, 3, 2, 1 till the 3 consecutive items have been passed or no items remain - whichever is first)
Record -subject must point to selection and examiner records the position # of that item
Score - the # of items answered correctly

9) Auditory-Decoding Test

Ceiling - A.If subject MA is 6-0 and passes 3 or more of the first 4 items, continue test till ceiling (when 4 in any 8 consecutive items are passed). If 2 or less of these first 4 items are passed - stop test.
B.If subject MA 6 or over, begin with item 5. If subject passes 3 or more of the 1st 8 items, continue to test till ceiling (when 4 in any 8 consecutive items are passed) and credit items 1-4. If subject passes 4 or less of items 5-12, administer items 1-4 and stop test.
Basal - Variable, as explained above.
Record - The subjects yes or no response (do not use +, -, or ✓.
Score - 1 pt. for each item correctly responded to.

TAKE____________ TAKE____________ TAKE____________ TAKE____________

TIME____________ TIME____________ TIME____________ TIME____________

MILWAUKEE TEACHER AIDE PROJECT MOTION PICTURE FILMING OPERATIONAL LOG SHEET

SCENE NO.__________ STORYBOARD NO.__________ SETTING BLOCK__________

PROJECT____________________ SCENE IDENTIFICATION____________________

AIDE: CURRICULAR AREAS:

DATE____________________ FILM____________ ASA__________ F-STOP______

BLOCK____________ PERIOD__________ CLASS____________ ROLL____________

NATURAL LIGHTING:

ARTIFICIAL LIGHTING:

NOISE LEVEL TEST, AMB.____________dB PEAK__________dB OFF______dbm

TEACHER__________dB AIDE__________dB CURRICULUM:

BAROMETRIC PRESSURE____________ Rising Steady Falling TEMPERATURE, IN________OUT________

WIND OUT OF__________AT__________MPH HUMIDITY______% ____________________

EXTENUATING CIRCUMSTANCES:

TECHNICAL:

POST [illegible]:

NOISE LEVEL TEST READS AN AVERAGE OF 5 dB LOWER WHEN SENNHEISER MD421 MICROPHONE IS USED INSTEAD OF SENNHEISER MKH 804 MICROPHONE.

TEST WITH 421 804 MICROPHONE

EMPLOYEE CHARACTERISTICS

	EXCELLENT	GOOD	FAIR	POOR
PERSONAL APPEARANCE				
PHYSICAL HEALTH				
EMOTIONAL STABILITY				
INITIATIVE				
ABILITY TO WORK WITH OTHERS				
RELIABILITY				
SYMPATHETIC UNDERSTANDING OF PROBLEMS				
ABILITY TO CONTROL				
SUB-PROFESSIONAL GROWTH				
COMMUNITY RELATIONSHIPS				
PUNCTUALITY				

Good natured & Agreeable ooooooooo oooo oooooo oooooooo(1)oooooooo oooooo oooo ooooooooo Stubborn & inclined to be difficult
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Intelligent ooooooooo oooo oooooo oooooooo(2)oooooooo oooooo oooo ooooooooo Unintelligent
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Tense & Nervous ooooooooo oooo oooooo oooooooo(3)oooooooo oooooo oooo ooooooooo Relaxed & Calm
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Physically Strong ooooooooo oooo oooooo oooooooo(4)oooooooo oooooo oooo ooooooooo Physically Weak
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Behaves Childishly ooooooooo oooo oooooo oooooooo(5)oooooooo oooooo oooo ooooooooo Behaves Maturely
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Old-Fashioned ooooooooo oooo oooooo oooooooo(6)oooooooo oooooo oooo ooooooooo Modern
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Bossy and Dominating ooooooooo oooo oooooo oooooooo(7)oooooooo oooooo oooo ooooooooo Submissive; tends to give in easily
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Thin ooooooooo oooo oooooo oooooooo(8)oooooooo oooooo oooo ooooooooo Fat
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Adventurous & Ready to Try Anything ooooooooo oooo oooooo oooooooo(9)oooooooo oooooo oooo ooooooooo Cautious and afraid of new situations
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Lazy ooooooooo oooo oooooo oooooooo(10)oooooooo oooooo oooo ooooooooo Ambitious
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Optimistic (Looks at the bright side) ooooooooo oooo oooooo oooooooo(11)oooooooo oooooo oooo ooooooooo Pessimistic (Looks at the dark side)
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Masculine ooooooooo oooo oooooo oooooooo(12)oooooooo oooooo oooo ooooooooo Feminine
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Young ooooooooo oooo oooooo oooooooo(13)oooooooo oooooo oooo ooooooooo Old
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Responsible & Determined ooooooooo oooo oooooo oooooooo(14)oooooooo oooooo oooo ooooooooo Irresponsible & Quitting
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Crude & Vulgar ooooooooo oooo oooooo oooooooo(15)oooooooo oooooo oooo ooooooooo Refined
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Tall ooooooooo oooo oooooo oooooooo(16)oooooooo oooooo oooo ooooooooo Short
Extremely Very Fairly Slightly Slightly Fairly Very Extremely

Continued - Page 2

Suspicious of Others	Extremely	Very	Fairly	Slightly	(17)	Slightly	Fairly	Very	Extremely	Trusting of Others
Thrifty	Extremely	Very	Fairly	Slightly	(18)	Slightly	Fairly	Very	Extremely	Wasteful
Dependent on Others	Extremely	Very	Fairly	Slightly	(19)	Slightly	Fairly	Very	Extremely	Self-Reliant
Cold-Hearted & Unsympathetic	Extremely	Very	Fairly	Slightly	(20)	Slightly	Fairly	Very	Extremely	Warm-Hearted & Sympathetic
Good Looking	Extremely	Very	Fairly	Slightly	(21)	Slightly	Fairly	Very	Extremely	Plain
Conventional (Tries to be Like Others)	Extremely	Very	Fairly	Slightly	(22)	Slightly	Fairly	Very	Extremely	Unconventional (Tries to be Different from Others)
Rich (in Money or Possessions)	Extremely	Very	Fairly	Slightly	(23)	Slightly	Fairly	Very	Extremely	Poor (in Money & Possessions)

Expectancy Index

Subject Code________

Class______________

Cell #_____________

Student Name:________________________

Rater Name:_________________________

On the basis of your knowledge of the above student, please indicate the extent to which you feel that he (she) will show marked improvement during the coming year in the indicated areas of functioning. That is, each child should be rated in each area from 0 (no chance of marked improvement) to 100 (certainty that he will show marked improvement). Consider each child separately and put a check mark above that number on the continuum which reflects your estimate of the likelihood that the child will show marked improvement.

Social Behavior (class adjustment):

0 10 20 30 40 50 60 70 80 90 100

Language (self-expression written and verbal):

0 10 20 30 40 50 60 70 80 90 100

Academics (subject matter skills):

0 10 20 30 40 50 60 70 80 90 100

Comments:

TAP Project

Second-Level Analysis

Name of Subject____________________ Level: Primary_______ Intermediate_______

Name of Analyst____________________ Date__________________

I. Type #1 communication units

Type	Run 1	Run 2	Run 3
a.			
b.			
c.			
d.			
e.			
f.			
g.			

II. Type #3 communication units

	Run 1	Run 2	Run 3
Total Number of Type #3 Units			
Total Number of Structural Complexity Points			
Mean Complexity Score			
Total Number of Well-Formed Type #3 Units			

Code No. ______
Cell No. ______
Test No. ______
ITPA Code______

Teacher-Aide Project

Speech Sample Transcription Analysis

Raw Indices

Student________________________Identification of Speech Sample______________

Analyst________________________Date________________

Number	Description	Score
1	Type 3 communication unit mean level of complexity . . .	______
2	Number of well-formed type 3 communication units	______
3	Time in minutes and five sec. intervals	______
4	Number of corrective segments	______
5	Number of communication mazes	______
6	Number of unresolved mazes	______
7	Total number of tokens in communication units	______
8	Total number of types in communication units	______
9	Unacceptable selected phonological tokens	______
10	Unacceptable morphological tokens	______
11	Unacceptable usage tokens	______
12	Unacceptable English word tokens	______
13	Number of syllables in transcript	______
14	Number of syllables in mazes	______
15	Number of finite verbs, including instances where a verb should be but was omitted	______
16	Number of incorrect verb forms (including omission of "to be") . . .	______
17	Number of omissions of "to be" with main verb following	______
18	Number of omissions of "to be" when required (i.e. no verb in sentence)	______
19	Number of type 1 communication units	______
20	Number of type 2 communication units	______
21	Number of type 3 communication units	______

Part II

Student____________________Code No.__________ Identification of Speech Sample__________

Analyst____________________Date______________

B. Derived Indices

Number		Formula	Score
1	Words per minute	6/1	______
2	Mean words per narrative segment	6/2	______
3	Syllables in mazes as a percent of total syllables in transcript	(3/8)100	______
4	Type-token ratio	(6/5)100	______
5	"Acceptable" words as percent of total uttered in narrative segment	[(5-7)/5]100	______
6	Unresolved mazes as percent of total number of mazes	(4/3)100	______
7	Percentage of verb forms incorrect	(11/10)100	______
8	Percentage of incorrect verb forms which are omission of "to be" with main verb following	(12/10)100	______
9	Percentage of incorrect verb forms which are omission of "to be": no verb in sentence	(13/10)100	______
10	Percentage of tokens more than one syllable	[(8-9)/5]100	______

Average number of syllables per token

Teacher-Aide Project

Speech Sample Transcription Analysis

Part I

Student__________________ Code No.____________ Identification of Speech Sample__________

Analyst__________________ Date______________

A. Raw Indices

Number	Description	Index
1	Time in minutes and five sec. intervals	______
2	Number of narrative segments	______
3	Number of communication units	______
4	Number of communication mazes	______
5	Number of unresolved mazes.	______
6	Total number of tokens in communication units	______
7	Total number of types in communication units	______
8	Unacceptable selected phonological tokens	______
9	Unacceptable morphological tokens	______
10	Unacceptable usage tokens	______
11	Unacceptable English word tokens	______
12	Number of syllables in transcript	______
13	Number of syllables in mazes	______
14	Number of finite verbs, including instances where a verb should be but was omitted	______
15	Number of incorrect verb forms (including omission of "to be") . . .	______
16	Number of omissions of "to be" with main verb following	______
17	Number of omissions of "to be" when required (i.e. no verb in sentence)	______

				T_1	T_2	T_3	
Primary	Aides	Stan	Bright				
			Meredith				
			Ahmadi				
		Stan, 95%	Moore				
			Conway				
	No Aides	Stan	Buchanan				
			Patton				
			Pittman				
		Stan, 95%	Brady				
Intermediate	Aides	Stan	Allen				
			Roberts				
			Hodges				
			Jones				
		Stan, 95%	Brady				
			Davis				
			Mann				
	No Aides	Stan	Beard				
			Richardson				
			Allen				
		Stan, 95%	Godwin				
			Gray				

Design of Study

		Aides		No Aides	
		Standard	Standard +50	Standard	Standard +50
Primary	Restrictions				
	No Restrictions				
Intermediate	Restrictions				
	No Restrictions				

DEPENDENT VARIABLE ____________________

			T_1	T_2	T_3
PRIMARY A_1	Allies B_1	Stan. C_1			
		Stan. + 50% C_2			
	No Allies B_2	Stan. C_1			
		Stan. + 50% C_2			
INTERMEDIATE A_2	Allies B_1	Stan. C_1			
		Stan. + 50% C_2			
	No Allies B_2	Stan. C_1			
		Stan. + 50% C_2			

SCHOOL ______________________

TEACHER ______________________

STUDENT	BIRTH DATE										
1.											
2.											
3.											
4.											
5.											
6.											
7.											
8.											
9.											
10.											
11.											
12.											
13.											
14.											
15.											
16.											
17.											
18.											
19.											
20.											
21.											
22.											
23.											
24.											
25.											

IDH - Oct. 1967

SCHOOL________________________

TEACHER________________________

Student	Birth Date														

		[illegible]				[illegible]				[illegible]				[illegible]				[illegible]				[illegible]		
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4

APPENDIX E

Analytical Procedures

1. Automatic Interaction Detector (AID)
2. Some General Comments on the Relation of Multivariate Analysis to Repeated Measures Analysis of Variance

1. Automatic Interaction Detector (AID)

The following discussion of AID, with minor changes, was taken from the book *Economic Behavior of the Affluent* by Barlow, R., Brazer, H.E., and Morgan, J.N.

An interaction effect exists where the explanatory or "casual" variables do not act independently and additively. The effect of one factor may depend on the level of another. Sometimes a variable affects only one part of the population. Sometimes several explanations are alternatives, as when *any one* of two or three circumstances leads to a low level of investment information.

The new program imposes no restrictions on the data, except that the dependent variable--the thing to be explained--must be a reasonably normally distributed variable or a dichotomy where the proportions do not get too close to zero or to 100 percent. It is really an automated replication of the process researchers have always used in looking for structural relations in a rich set of data with many variables and a large number of cases.

The parallels between the program and the strategy of an exploratory researcher are striking. More often than not, investigators have only a rudimentary theory and must generally let the data speak for themselves. Even if they first subordinate the data to a tight test of a prior hypothesis, they generally proceed to a great deal of secondary searching or revised hypotheses. They often fix attention on a single dependent variable, the aim being to find independent variables whose categories separate the sample into widely differing subgroups, each internally homogeneous. This feature underlies regression models as well as analysis of variance models. But in a less formal way researchers often conduct an intuitive and unsystematic search for homogeneous subgroups without conceptualizing it as such. One of the program's virtues is that it bares this approach. Using it, the researcher must face squarely the exploratory nature of his method. Because it is made explicit, the method is "exposed to fraternal judgment." AID merely formalizes a common strategy and gives it astronomically greater power by harnessing the computer to it.

The program operates sequentially. It first looks at the effect on the whole sample of each explanatory characteristic in turn, finding the best way of using that factor to divide the sample into two groups. By "best," we mean the division that reduces the unexplained sum of squared deviations (or "error variance") the most. Using two subgroup averages instead of one overall average to predict the dependent variable reduces the predictive error, and the process looks for the largest such reduction. The test is a test of importance, not significance. The two groups so formed must not only be different, they must also each contain a substantial number of cases. If a few cases can be split off and the unexplained variance reduced substantially, the dependent variable clearly contains extreme cases (is heteroscedastic), a highly dangerous situation for any statistical analysis based on least squares. Within each of the two groups, the individuals are more like one another (homogeneous) than they are like those in the other group.

Where the explanatory characteristic has a natural ordering or rank order, that order can be preserved, but it is also possible to try for the best re-ranking and then the best division of the sample along that new ordering.

The program then recalls which predictor's best split is better than that of the other predictors, and actually divides the sample into two subgroups on that basis. Each of the two subgroups thus generated is examined in turn in the same way, using all the predictors, and is split again. Each of these new groups is examined in turn and split if possible. The new groups spread out like the roots of a tree, containing fewer cases, each split doing its best to maximize the explained variance (minimize the unexplained variance).

The process stops when no division of a group can reduce the unexplained sum of squares (variance) by as much as some predetermined criterion level, usually 0.5 percent or 1 percent of the original full-sample sum of squares (around the average). Or it stops when one or both of the groups produced would be so small that there would be a serious question as to the sampling stability of the results--that is, we could not be reasonably sure that another sample would show anything similar. Some of the roots end before others, of course, and different predictors are used in different roots.

Since the search process looks at so many possibilities, it cannot be thought of as testing a hypothesis. There are no "degrees of freedom" left and no way to estimate the stability of the results over different samples. Tests with the program indicate that the first parts of the process remain reasonably stable with different samples, and the groups developed somewhat more stable than the path by which they are generated. This problem exists whenever a number of analyses are made of the same body of data, but the present formalization of the process makes it more obvious.

There is one way in which we can come to a conclusion about the effect of a particular predictor. If it has no apparent effect either over the whole sample, or on any of the major subgroups generated, then one can be reasonably sure it does not matter. (The program prints out the subgroup averages according to each predictor at each stage.)

Even this certainty is reduced where there are two predictors that are correlated with one another. With the sequential approach, once the sample is divided on the basis of the more important predictor, the other is often unable to assume any importance at all. With multiple regression, by comparison, both predictors would show some apparent effect, dividing up the credit, but the sampling errors tend to become large.

The results of the process are independent of the order in which the variables are introduced, but they do depend on which variables are allowed into the analysis, and on the precision with which they are measured. It is always possible that introducing a different variable or a better measurement of one already used, would produce different results.

The program should not be confused with the large number of multivariate analysis programs for data reduction now in use based on multiple regression, multiple discriminant analysis, or factor analysis. All these other programs assume additivity of effects of different factors in one way or another. They impose various other restrictions to make the problem manageable. In the case of factor analysis, the factors are sets of weights (loadings) assigned to explanatory variables and used to provide a new, smaller set of indexes used as predictors. The weights are based on intercorrelations among the variables, however, not on their relations to the variable to be explained.

The new program is, of course, no substitute for theory, which is involved in the selection of the variables and the interpretation of the results. Nor does it provide the final word, since ultimately the findings need to be tested against other samples and at a later date for stability and persistence. What has been done here is to focus on the factors that seem to matter now.

In a few cases where it was important to assess the effect of one predictor relatively free of spurious effects through other variables, and where a substantial amount of symmetry (additivity) seemed to exist, a multiple regression with dummy variables was used.

On the other hand, previous tests have shown that frequently a predictor that had a significant multiple regression coefficient turned out in the new analysis to be affecting only one subgroup of the population.

2. Some General Comments on the Relation of Multivariate Analysis to Repeated Measures Analysis of Variance

In the present study, for each response variable such as reading achievement there were three responses for each individual taken at three points in time.

In both analytical approaches the objective is to determine if the three responses have a common expected average value or whether the three have different average expected values. In addition, in both analyses one attempts to determine if differences between the three measures vary from group to group or condition to condition, i.e., whether the amount of change over time is the same or different for all groups.

In the repeated measures analysis one assumes that the three responses are equally correlated with each other and have equal variances. Thus, orthogonal contrasts between them (there are two linearly independent orthogonal contrasts) would be statistically independent and have the same variance. Therefore, each pair of contrasts would provide two observations which could be analyzed by ordinary fixed effect analysis of variance.

The multivariate analysis treats the general case when the three responses are not equally correlated or do not have equal variances. Again, two contrasts (two linearly independent but not necessarily orthogonal contrasts) are used as two dependent variables which are treated as though they would be correlated.

Both of these new variables reflect or would indicate differences over time as they are differences between the three responses of each individual. One might do a separate univariate analysis of variance on each one but the two F-tests for each effect would not be independent. Also, and perhaps more importantly, each would be incomplete as one would want to infer change or lack of change from both differences rather than just one. The multivariate analysis of variance can be thought of in the same way as a univariate analysis of variance. In the former the dependent variable is a two dimensional variable, e.g., change has two dimensions, say change from period one to the average of periods two and three and change from period two to three (of course, there are other ways of viewing the two dimensions of change). The multivariate ANOVA does not require the assumptions of equal correlations among the original variables.

The two-stage procedure

In the first stage the hypothesis of no change on any type of score (such as reading, arithmetic, and spelling as obtained from the WRAT) against the hypothesis of some change on at least one. [Or, the null hypothesis of no difference between conditions of change of any type against some difference of some type between conditions.]

If the first stage rejects the null hypothesis, we then try to determine if change occurred on all types of scores, just two, just one, or indeterminate. The three analyses were done using reading scores, arithmetic scores, and spelling scores. To control the type I error and make it .05 for all three analyses, we require significance at the .05/3 level. That is, if we just did the three analyses and used a .05 significance level on each, the probability that we would *falsely* reject the null hypothesis for the three analyses together would be greater than .05.

Rejection of the null hypothesis at stage I says there has been some change of some type (or there is difference between conditions of some type) while lack of rejection says no change of any type or no difference between conditions of any type. At stage II we try to pinpoint on which type of achievement the change or difference occurred without greatly increasing the type I error rate.

3. Modified WRAT Administration Instructions

Because of the somewhat unique nature of pupils involved in the study and the fact that we had to train our own people (aides and graduate students) to administer the WRAT, a somewhat modified, albeit consistent, administration procedures were developed. These modified procedures were as follows.

In general, this is a summary of the instructions contained on pages 3-4 of the manual.

Before administering the test, the examiner is encouraged to see that the following blanks of the personal data lines are filled in: Name, birth date, school, date, and examiner. Also circle M or F to indicate the sex of the subject.

All subjects are to be given the Level I test regardless of age (instructions in the manual not withstanding).

Begin the Spelling Test with Test Instructions--Spelling Level I, found on page 3 of the manual. For Item #1, Copying Marks, read the instructions to the testee as given. Permit 60 seconds for entire row, as indicated. For #2, Name, read the instructions as given, again permitting 60 seconds for this part. (Since for the Reading test, part of the test is to ask the testee to name the first two letters of his name, this may be done at this time; as per #3 under Pre-Reading Section on page 3) For Dictation of Words (Item #3, Spelling Test), read instructions to the testee as given. Words for this test are found on page 12. Allow 15 seconds for each word, and continue dictating until the testee has missed 10 words in succession. The testee should be instructed to draw a line in the space if he does not know how to spell a word.

Begin the Reading Test with the Pre-Reading section, instructions for which are given on page 3 of the manual. Again, this section should be given exactly as directed. For the Pre-Reading section, responses should be recorded by the examiner by underlining correctly named or recognized letters and/or by crossing out incorrectly named or recognized letters.

Upon completion of the Pre-Reading section, give the Reading section as per directions on page 4.

Begin the Arithmetic Test with the Test Instructions at the bottom of the second column on page 3. If the testee responds correctly to the first problem, continue the instructions and permit testee to begin the computations. While testee is working, note whether or not he is getting correct answers. If he does not get at least five correct answers on the first line, question him as to whether or not he feels he is going to be able to do the rest of the problems. On the basis of the response to this questioning, make a determination as to whether or not it would seem practical to allow him to continue with the written part. If it is determined that it would not be practical, go immediately to the oral part of the arithmetic test, instructions for which are given in the first column of page 4, beginning with #1, Counting Dots. If the testee does get at least five answers correct on the first line of the written test, permit him to continue working for the remainder of the ten minutes.

Testees who are given 10 minutes for the written test are to be given the oral portion of the test at the completion of this time, regardless of their score (instructions in the manual again not withstanding).

The oral part of the test should be administered exactly as per instructions in the manual. Time limits and instructions for recording responses, both of which are found on page 6, should be observed and carried out.

In timing the test, at least for those portions for which the time limits are very short, it is not necessary to start and stop the stop-watch for each item, but merely to note the time at the beginning of the item and what it should be when the time for the item is up.

3. FORMS USED IN STUDY

CLASS SIZE ____________

INCREASED ____________

STANDARD ____________

TEACHER QUESTIONNAIRE
CLASS SIZE AND TEACHER AIDES IN EMR CLASSES
JUNE, 1967

Please place a check following each item to indicate your judgement of the situation in your classroom during this semester in comparison to last semester.

	MUCH BETTER	SOMEWHAT BETTER	NO CHANGE	NOT AS GOOD	NOT NEARLY AS GOOD
PUPIL OUTCOMES:					
1. Pupil Achievement					
2. Expressive Verbal behavior a. Oral					
b. Written					
3. Pupil Interest in Learning					
4. Initiative in Academic Work					
5. Behavior in General					
6. Social Attitudes toward Peers					
7. Attitude toward School					
8. Pupil Attendance					
TEACHER ACTIVITY					
1. Individualization of Instruction					
2. Flexibility a. Instructional Arrangement					
b. Variation in Content					
c. Variation in Method					
3. Teacher feeling of Accomplishment					
4. Teacher feeling of Satisfaction					

5. Please comment on the effectiveness of the Peabody Language Development Kit.

Advantages: ____________

Disadvantages: ____________

Overall Usefulness: ____________

[illegible] Would you like the same instructional arrangement in your classroom next year?

Yes ______ No______ If not, what changes would you suggest? ____________

[illegible]

NAME ____________________
CLASS ____________________
SCHOOL ____________________

______This teacher aide ______________________________has many desirable traits and through her services to the classroom teacher, she contributes much to the children as well as to the school program. The success of the research project is, in large measure, due to teacher aides that possesses and exercises a well-rounded combination of abilities. This teacher aide is a contributor to this success. Therefore, it is requested that she continues in her building performing the same role.

____

____

____

____

____

________In the opinion of the On-Site Coordinator, ____________________________, this teacher aide is not making a satisfactory contribution to the research program. Whil she possesses certain good qualities, nevertheless it is felt that she fails to meet the standards necessary for the successful operation of the Teacher Aide Research Project. Under these circumstances, it is recommended that she not be included in the project for the next semester. The following explicit justification for the recommendation is given below.

____

____

____

____

____

TAP - Film Sheet

Cannister ID____________________ Date of Shooting________________

First Half______________________ Second Half_____________________

Primary________Intermediate________; Aide________ No Aide________

Standard Size___________________ Increased Size__________________

I. Photography

1. Focus: Blurred ____ ____ ____ ____ ____ Clear

2. Panning: Erratic ____ ____ ____ ____ ____ Smooth

3. Zooming: Erratic and Blurred ____ ____ ____ ____ ____ Smooth and Clear

4. Lighting: Bad ____ ____ ____ ____ ____ Good

II. Sound

1. Intelligibility: Poor ____ ____ ____ ____ ____ Excellent

2. Volume: Variable ____ ____ ____ ____ ____ Steady

III. Appeal

Humorous: Little ____ ____ ____ ____ ____ Much

General Interest: Little ____ ____ ____ ____ ____ Much

Activity Level: Little ____ ____ ____ ____ ____ Much

Distracted by Camera: Little ____ ____ ____ ____ ____ Much

IV. Content

1. Curriculum Topic______________________________________

2. Aide Activity____

3. Teacher Activity_______________________________________

4. Grouping One Large__________________Several Small__________

5. Other Observations_____________________________________

____

DATE ____________________

TEACHER'S NAME ____

SCHOOL ____

TYPE OF CLASS ____

TAKE NUMBER ____

In order to do a more comprehensive analysis of the film sequences which we are taking in your classroom, it is imperative that we have information concerning the activities represented in these sequences. This information will be an aid for an understanding of the context of the classroom activities for interpretation by the research analysts.

The basic information we need concerning the sequences are:

1. If the children are grouped into small groups, why are they grouped as they are?

   ____

   ____

   ____

2. If the children are all in one group, what is the purpose of this grouping?

   ____

   ____

   ____

3. What curricular area is each group working in?

   ____

   ____

   ____

4. What specific lesson or concept is being worked on?

   ____

   ____

   ____

5. Where does this lesson fit into the broader curricular area?

6. If you have an aide, what is the role of the aide in this sequence?

7. Note any prevailing circumstances at time of filming, for example, just before a holiday, before or after recess, etc.

You may be very brief in answering these questions. Long explanations are not necessary.

Thank you for your cooperation. This information will contribute to a more successful research study and comprehensive research document.

By______________________
(Mrs.) Sarah Scott
On-Site Coordinator

Meteorological and Noise Indices

Teacher-Aide Project

Class:________________________

		Noise (db)				Meteorological						
Block	Period	Amb.	Peak	Teacher	Aide	Bar. Press.	Dir.	Temp in	Temp out	Wind	Humidity	Wind Dir.

School____________________

Teacher___________________

Teacher-Aide Project

Absentee Form

Student Name	Semester	Date Entered Class	# of Days Should Have Been in School	# of Days Absent from School	Absentee Date
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				
	Spring, 1967				
	Fall, 1967				
	Spring, 1968				

TEACHER ACTIVITY LOG

School ______________________ ☐ Primary ☐ Intermediate

To the teacher

In order to help us learn how teacher aides are assisting us in our school work with children, we are asking you to record your activity on a single day during a month. We would like you to tell us the time you started and finished each activity during the day. Please be quite specific and as objective as possible.

This information is for research purposes only and will in no way be related to an evaluation or rating of an individual person.

Please mail to Mrs. Scott at the end of the day.

TIME From:	To:		ACTIVITY
Example A 8:45	9:00	A	Conference with a student's mother regarding his reading ability
Example B 9:00	9:05	B	Planned for arithmetic lessons for this morning

TIME From:	To:		ACTIVITY
No. 1		No. 1	
No. 2		No. 2	
No. 3		No. 3	
No. 4		No. 4	
No. 5		No. 5	

TIME ACTIVITY

From: To:

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

No. ____ __________ __________ No. ____ ____________________
____

TIME ACTIVITY

From: To:

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

No. ___ ____________ ____________ No. ___ ______________________
____

Subject Code ____________

Class ____________

Cell Number ____________

Behavior Rating Scale

Name ____________ Date ____________

Sex ____________ Teacher ____________

Birthdate ____________ School ____________

Rating Guide

1. Base rating on pupil's behavior since the beginning (end) of the semester.
2. Base your rating on your own experience with the pupil being rated.
3. Complete each item without consideration of ratings on other items.
4. Base your ratings on observations of outward behavior and not on what you may think may underlie this behavior.
5. Do not spend more than a few seconds on any item. Rate the item quickly.
6. Complete all items.
7. Each item is to be rated with a 1, 2, 3, 4, or 5 according to the following:

   1 = Does not describe; Never

   2 = Seldom; Rarely

   3 = Occasionally; Somewhat descriptive

   4 = Often; More than usual

   5 = Very frequently; Describes to considerable extent, almost always

Behavior Rating Scale

Check One

Item	Does not describe; Never 1	Seldom; Rarely 2	Occasionally; Somewhat descriptive 3	Often; More than usual 4	Very frequently; Describes to considerable extent 5
1. is able to concentrate on things					
2. is disruptive and boisterous					
3. if passed over or slighted or if things go wrong, is inclined to sulk					
4. is interested in schoolwork					
5. is alert in class					
6. demands the attention of the teacher and pupils					
7. displays feelings of inferiority					
8. is popular with classmates					
9. threatens to hurt other people when angry					
10. is stubborn					
11. companions find it difficult to get along with him					
12. becomes embarrassed easily					
13. is timid in meeting people					
14. uses profanity					
15. finishes his classroom assignments					
16. has changeable moods					
17. resents even the most gentle criticism of his work					

Check One

Item	Does not describe; Never 1	Seldom; Rarely 2	Occasionally; Somewhat descriptive 3	Often; More than usual 4	Very frequently; Describes to considerable extent 5
18. seems to be off in a world of his own					
19. maintains same facial expression (blank stare or far-away look					
20. is uncooperative					
21. is tired and listless - little energy					
22. when angry will refuse to speak to anyone					
23. seems to be drowsy					
24. gives an impression of irresponsibility					
25. can be trusted with responsibility					
26. displays little interest in his environment					
27. appears to be preoccupied					
28. shows a dislike for school					
29. daydreams					
30. gives other pupils dirty looks					
31. has to have everything his own way					
32. works well by himself					
33. tries to arouse the sympathy of others					
34. uses abusive language toward others					
35. gives up when faced with difficulty					

Check One

Item	Does not describe; Never 1	Seldom; Rarely 2	Occasionally; Somewhat descriptive 3	Often; More than usual 4	Very frequently; Describes to considerable extent 5
36. feels shy when the center of attention					
37. has a hard time making decisions					
38. shows little interest in classroom routine					
39. is afraid of strange adults					
40. seeks affection from others					
41. postpones own work to help others even though does not have time to do his own					
42. "drags his feet" when requested to do something					
43. seems unconcerned when he misbehaves					
44. does not forget things that anger him					
45. does not participate in group activities, stays in background					
46. becomes frightened easily					
47. can be trusted to carry out instructions promptly and attentively					
48. gives the impression of being lazy					
49. is discontented with everything in general					
50. prefers to be alone - works and plays alone					
51. finds fault with instructions given by adults					
52. acts up when I'm not watching					

Check One

Item	Does not describe; Never 1	Seldom; Rarely 2	Occasionally; Somewhat descriptive 3	Often; More than usual 4	Very frequently; Describes to considerable extent 5
53. has difficulty disagreeing with others as he fears disapproval					
54. does not respect other people's belongings					
55. seeks out adult help in doing things					
56. is a helpful person					
57. appears to be sad, unhappy, lonely					
58. when criticized by a teacher frequently reacts in a sullen manner - sulks or pouts					
59. readily becomes enthusiastic over a new project or activity					
60. has a "chip on his shoulder"					
61. has difficulty in controlling temper - is easily upset					
62. seeks approval for his behavior					
63. acts dependent					
64. lacks self-confidence					
65. usually complies when others try to boss					
66. seeks adult attention					
67. is inactive or sluggish					
68. volunteers to help when there is work to be done					
69. is considerate of others					
70. is upset by criticism or disapproval					